ON THE MEASUREMENT OF COLOR AND BRIGHTNESS

CEES H. ELZINGA

ISBN 90-9000878-0

ON THE MEASUREMENT OF COLOR AND BRIGHTNESS

Proefschrift

ter verkrijging van de graad van doctor
in de sociale wetenschappen
aan de Katholieke Universiteit van Nijmegen
op gezag van de rector magnificus
Prof. Dr. J.H.G.I. Giesbers
volgens besluit van het College van Dekanen
in het openbaar te verdedigen op
vrijdag 15 maart 1985,
des namiddags te 2 uur precies

door

Cornelis Hessel Elzinga
geboren te Hilversum

Promotor: Prof. Dr. E.E.Ch.I. Roskam

Co-referent: Dr. Ch.M.M. de Weert

Preface

This study was supported by a grant from The Netherlands Organization for the Advancement of Pure Research (ZWO), project number 15-31-10.
During the final stage of preparing and writing this booklet I was greatly inspired, supported and accurately assisted by my dear friend Marty Rohrbach.
For the design and maintenance of the equipment, as well as for software, I, sometimes strongly, depended upon the cooperation of the technical staff of the laboratory. Without its expertise and efficient service this work could not have been realized.
Of course, nothing could have been accomplished had there not been many colleagues with whom it was a pleasure to work.

Contents

Chapter 1:

THEORIES ON COLOR AND BRIGHTNESS, MEASUREMENT AND MEANINGFULNESS

Introduction

For centuries hypotheses on the structure and function of the visual system have been proposed to explain our everyday sensations of brightness and color. Since Newton, interest has focussed on the determination of the characteristics of the receptors of the human eye and on the way the interacting outputs of these receptors trigger the various sensations.

In order that quantitative relationships between receptor-outputs and sensations of color and brightness may be established, these outputs and sensations themselves need quantification. The measurement of the receptor-outputs is largely a problem of physics although their actual assessment has originally been accomplished through the use of psychophysical methods.

In this chapter we will discuss a number of questions that concern the measurement of sensations and the specific problems that arise in connection with color and brightness. Our attention will focus on theories that permit of a measurement structure in which both physical and psychological variables occur.

As soon as one attempts to measure sensations problems arise that, at first sight, appear to be very different from measurement problems in physics. On the other hand, measurement procedures that are, formally, identical to procedures used in physics, are frequently used to measure sensations.

A well-known example is the measurement of the intensity of the sensations of redness and greenness and of yellowness and blueness as first performed by Hurvich and Jameson (1957). They measured these sensation intensities through

a cancellation technique that is very similar to the procedure of measuring acidity: determine the power (concentration) of a reddish light (an acid) that, when admixed to a greennish light (an alkaloid) produces a light (a solution) that neither has the quality of redness nor the quality of greennes (contains an equal number of hydrogen and hydroxide ions).

Appealing as such a procedure may seem, it is, however, not immediately clear to what the numbers refer that are the result of such a procedure. Are they indeed to be considered as reflecting the intensities of the sensations of redness and greenness and, if so, do they really imply more than an ordering of these intensities?

As a second example, consider the following two statements:

(i) "this photodiode is five times as responsive to 1 Watt/cm^2 at 560 nm than it is to 1 Watt/cm^2 at 460 nm"

(ii) "the brightness of light of 1 Watt/cm^2 at 560 nm is five times that of light of 1 Watt/cm^2 at 460 nm".

Many people probably will wrinkle their eyebrows at reading the second sentence since it implicitly refers to the existence and establishment of units of brightness, just like the first sentence refers to units of electrical power.

Below we will therefore make some introductory remarks on the theory of measurement and meaningfulness in connection with theories on color and brightness.

Omitting spatial and temporal variations, incoherent radiation is physically specified by the amount of radiant power for each of its wavelengths. To each bundle of radiation a we may then assign a function $D_a(\lambda)$ which provides the radiance density at each wavelength λ.

The physical structure, or rather the extensive measurement structure for radiant energy, then permits us to also compute such a function for incoherently superimposed bundles a and b as

$$D_{a \odot b}(\lambda) = D_a(\lambda) + D_b(\lambda).$$

Here, the binary operation of incoherent superposition is indicated by the symbol $\odot$. Furthermore, this measurement structure determines such density functions to be homogeneous with respect to changes of the radiance level, indicated by the operation *, by some positive fraction t:

$$D_{t*a}(\lambda) = tD_a(\lambda).$$

Physically, this amounts to placing a neutral density filter between the light source and the measuring device, such that after passing the filter, the radiant power is reduced with a fraction t. Stated more formally, this simple physical structure allows for an isomorphic embedding onto a convex cone over the real numbers (Krantz, 1975a).

If the wavelength of the radiation is within the range of 400 to 700 nm and of sufficient power, it is visible to the normal, unimpaired human observer. The sensations this radiation evokes are of various kind and may either, like hue, be represented on a metathetic continuum or, like brightness or saturation, be ordered on a prothetic continuum.

If we consider the light bundles as stimuli, the sensations they arouse may be considered as attributes of these stimuli and, in doing so, we enrich the simple physical structure with an additional relational structure that is perceptual rather than physical. It is this enriched relational structure that is the proper subject of the psychophysics of color and brightness vision.

The most primitive aspects of vision probably are those of discrimination and detection. Hence, the bulk of research on vision, as well as in most other branches of psychophysics and psychophysiology, uses paradigms based upon various types of discrimination and detection experiments.

In the eighteenth century it was already noticed that bundles with distinctly different density functions can be made indistinguishable to the human observer, provided these density functions are suitably chosen. Since that discovery a huge proportion of the effort in vision research has been devoted to the careful specification of the density functions that are indiscriminable to the human eye if presented simultaneously and adjacently. This type of research, its results and the applications thereof are nowadays referred to as "colorimetry".

The main, outstanding result of colorimetry has been the establishment of mappings of lights in three-dimensional vector spaces over the real numbers, such that indiscriminable lights are mapped onto the same vector. This result may be considered as the first enrichment of the physical structure with a relational constraint that derives from perceptual data.

Since these mappings are unique up to non-singular linear transformations (Krantz, 1975a) and, furthermore, depend upon viewing conditions, several of such mappings, along with associated two-dimensional projections (chromaticity diagrams), have been accepted as standards by the Commission Internationale

d'Eclairage (CIE) (Wright, 1946; Wyszecki and Stiles, 1967; Estévez, 1979).

The fact that the dimensionality of the color space is limited has provoked explanations in terms of the physiology and anatomy of the retina. Although the hypothesis that the retina contains three different classes of receptors usually is credited to Thomas Young (1802), Walls (1956) and MacAdam (1970) consider Palmer (1777) as the originator of this idea.

Today, it is beyond much doubt that precisely three classes of receptors are mediating normal photopic vision. Brindley (1960a) amply describes the wealth of data that has led to this general agreement and discusses what little doubt remains. Recently, Barlow (1981, 1982) and Buchsbaum and Gottschalk (1983) also advocated this hypothesis on the basis of arguments of efficiency of information processing. This so-called "three pigment hypothesis" has become the starting point for most theories on (color) vision, the discrepancies among which arising from differences in paradigm and empirical phenomena that were chosen to append to the structure of the color-matching space.

Helmholtz (1896) aimed at a line-element model based upon color-discrimination, whereas Schrödinger (1920) took brightness-discrimination as the empirical point of departure. König and Dieterici (1892) started from the phenomena of color-blindness to reveal the fundamental cone-mechanisms whereas Wright (1946) relied upon phenomena of adaptation and Stiles (1946, 1959) used both adaptation and discrimination techniques. The way led by Hering (1878), Hurvich and Jameson (1957) employed the appearingly opponent structure of hue sensations as the basis of their quantification of many phenomena in color and brightness vision. Guth (e.g. Guth, Massof and Benzschawel, 1980) started from the nonadditivity of luminance to construct a theoretical framework that encompasses an abundance of data.

This listing of outstanding contributors to color science is by no means complete and the ideas proposed by any one of them certainly are worth elaborating upon. Comprehensive reviews are provided by, for example, Wasserman (1978) and Boynton (1979).

We will consider the question as to what kinds of perceptual data are relevant to the construction of psychophysical theory.

Class A and Class B observations

Given a century of building up theories and continuous growth of data, this question may seem obsolete. Unfortunately, this is not so since, when asked for, many colleagues in the field of vision research quite unequivocally and eloquently express beliefs on this matter that, as will be shown, are quite unfounded. This body of prejudicial attitudes towards the relevance of observations and the inferences possibly drawn from them is perhaps most clearly and explicitly stated by Brindley (1960a). He was the first to formally introduce the distinction between Class A and Class B observations and to discuss what he saw as their respective roles in theory construction and the testing of hypotheses on the visual pathway.

According to Brindley (1960a, page 133) Class A observations are direct reports of subjects on the (non-)identity or (non-)detectability of visual stimuli. Class B observations, on the other hand, are those not belonging to Class A and "... include all those in which the subject must describe the quality or intensity of his sensations or abstract from two different sensations some aspect in which they are alike". The appropriateness of Class B observations is then commented upon by his statement: "The most completely satisfactory way of using a Class B observation is to convert it into one

belonging to Class A".

Brindley himself admits that such a conversion is, in most instances of a Class B observation, impossible This position excludes numerous sensations from vision research that belong to everyone's daily experience with visual stimuli.

This problem is indeed recognized by some who adhere to this attitude. Teller and Pugh (1983), for example, attempt to formalize it but fail to give a full account of the logical status of paradigms prominent in vision research.

In order to explain our own views on these matters we will first make more explicit what exactly is meant by the hitherto loosely employed terms "relational structure" and "measurement structure" in order to use these formalizations to outline a concept of meaningfulness.

Measurement Structures

Loosely speaking, a measurement structure consists of a set, the elements of which are to be interpreted as real-world objects, and a (sub-)set of a real vectorspace such that the relations and operations on the real-world set are somehow reflected in corresponding relations and operations on the numerical set.

In this paragraph we will elaborate upon these concepts and provide formal definitions that make explicit what exactly is meant by "somehow reflected" This formalization heavily relies upon the concept of relation which will often be used in this and the next paragraphs. Since some readers may not be familiar with this concept, a short discussion of it may be of some help.

Let A be a nonempty set. We could interpret A as the set of all visible lights. We know that some lights appear brighter than others. If light a appears brighter than light b we could write. aBb It is, however, not true

that aBb for any pair of lights (a,b) so for some pairs the expression aBb may be correct while it is not for other pairs. Formally, any pair (a,b) with a,b$\in$A is an element of the set A×A: (a,b)$\in$A×A for all a,b$\in$A. We may now define the subset K$\subset$A×A of ordered pairs (a,b) such that (a,b)$\in$K if and only if aBb (in the sequel we will often use the word "iff" for the expression "if and only if"). The subset K is called a relation on A and it is a binary or 2-ary relation since its elements are ordered pairs.

As a second example, imagine a subject presented with the task to decide in which of the pairs of lights (a,b) and (c,d) the hue-difference is greater. Suppose this subject chooses the pair (a,b) as the one in which the hue-difference is greatest. We could then write: (a,b)H(c,d). Analoguous to the previous example we may now construct the Cartesian product set A×A×A×A with the subset L defined by (a,b,c,d)$\in$L iff (a,b)H(c,d). L is said to be a quaternary or 4-ary relation on A and any subset of A is a unitary relation on A.

Generally, a k-ary relation on a set A is a set of ordered k-tuples ($k\geq 1$), composed of elements of A.

Often relations may be characterized by stating their properties instead of listing their elements and, if so, may be given special names depending on the properties. Next we discuss a few elementary properties of binary relations that may be used to define various types of relations that are prominent in this capter and the next ones.

Let A be a set with elements a,b,c,... and R a binary relation on A. R is called "reflexive" iff (a,a)$\in$R for all a$\in$A. R is called "symmetrical" iff (a,b)$\in$R implies (b,a)$\in$R for all a,b$\in$A and R is called "transitive" iff (a,b),(b,c)$\in$R implies (a,c)$\in$R. Finally, R is called "connected" iff, for all

a,b∈A, either (a,b)∈R or (b,a)∈R.

The above properties may now be used to identify certain types of relations. First, a relation R is said to be an equivalence (relation) iff R is reflexive, symmetric and transitive. Second, a relation R is said to be a weak order (relation) iff R is reflexive, transitive and connected.

We are now in a position to explicitly define the concept of a measurement structure. The next three definitions capture the notions of relational and measurement structures and make explicit what exactly is meant by the expression "somehow reflected" that was used in the previous line. Similar definitions occur in the rich literature on measurement theory (Suppes and Zinnes, 1963; Pfanzagl, 1968; Krantz, Luce, Suppes and Tversky, 1971; Roberts, 1979).

Definition 1. Let Re denote the set of real numbers. The (n+1)-tuple $X=<A,R_1,\ldots,R_n>$ is a relational structure iff A is a non-empty set and for each i, R_i is a k_i-ary relation ($k_i \geq 1$) on A. The set A and the relations R_i are called the primitives of X. Let $Y=<B,S_1,\ldots,S_n>$ be another relational structure. Then X and Y are said to be of the same type iff, for each i, R_i and S_i are both k_i-ary. The relational structure X is called numerical iff $A \subseteq Re^m = Re \times Re^{m-1}$, $m \geq 1$.

Definition 2. Let $X=<A,R_1,\ldots,R_n>$ and $Y=<B,S_1,\ldots,S_n>$ be relational structures of the same type. Let ψ be a mapping from A into Y such that, for all i,

$$(a_1,..,a_k) \in R_i \quad \text{iff} \quad \{\psi(a_1),..,\psi(a_k)\} \in S_i$$

Then ψ is called a homomorphism from A into B, X is called the domain of

discourse and Y is called a representation of X. ψ is called an endomorphism iff A=B. ψ is called an automorphism iff it is an endomorphism that is one-to-one and onto B.

Definition 3. Let $X=<A,R_1,\ldots,R_n>$ and $Y=<B,S_1,\ldots,S_n>$ be relational structures of which Y is numerical. The triple $M=<X,Y,\Psi>$ is said to be a measurement structure iff the set of homomorphisms $\Psi=\{\psi|\psi.A\rightarrow B\}$ is not empty.

Definition 3 is not indisputable because some authors (Coombs, Raiffa and Thrall, 1954; Roberts, 1979) do not require Y to be numerical and R_1 is sometimes required to be a weak order (e.g. Luce, 1978).

Measurement theory is then concerned with three questions. The first is the problem of establishing, given X and Y, the non-emptiness of Ψ, the solution of which is to be formulated in a so-called representation theorem. The second problem is the description of the elements of Ψ, i.e. to describe what relations exist between different elements of Ψ. The result of such a description is then stated in the form of a so-called uniqueness theorem. The third, and certainly not the least problem is to provide a sound empirical denotation of the properties of the non-numerical structure because demanding a non-empty Ψ with a relatively simple structure quite often seems to necessitate a rather rich structure on the non-numerical set, involving properties of completeness, density and solvability whose correspondence with real-world properties is all but obvious.

Many representation theorems that have been proven concern a representation of which the primitives are a numerical set, an order or an equivalence relation and one or more operations like addition and multiplication. However, such

operations have equivalent relations: for example, the binary operation of multiplication on Re has an equivalent ternary relation S, defined by $(x,y,z)\in S \leftrightarrow x\cdot y=z$ for all $x,y,z\in Re$. More generally, let R be a ternary relation on a set A. Then R defines an equivalent binary operation on A iff, for all $a,b\in A$, there exists a $c\in A$ such that $(a,b,c)\in R$ and, for all $a,b,c,d\in A$, $(a,b,c)\in R$ and $(a,b,d)\in R$ imply $c=d$. Analoguously, to each k-ary operation there exists an equivalent (k+1)-ary relation on A. So, the previous definitions of measurement structures was stated in terms of relations only.

Krantz (1975a,b) was the first to provide a formal, measurement theoretical account of the density functions mentioned on page 3. Again, let A denote the set of lights and $\odot$ and * the binary operations of incoherent superposition of lights and intensity manipulation, respectively, and $X=\langle A,\odot,*\rangle$. Furthermore, let $Y=\langle Re,+,\cdot\rangle$ denote a vector space over the real numbers. It is a simple corrollary to theorems proven by Krantz (1975a, Theorems 1, 2 and 3) that $M=\langle X,Y,\Psi\rangle$ is a measurement structure such that for each $\psi\in\Psi$, all $a,b\in A$ and all $t,u\in Re$,

i) $\psi(t*a\odot u*b) = t\cdot\psi(a) + u\cdot\psi(b)$.

More important is, that Krantz (1975a) enriched this "physical" structure X with an additional relational constraint that derives from perceptual data. Let X denote the relational structure from the previous theorem and let us interpret A as the set of all visible lights. To incorporate the main results from colorimetry, Krantz defined an equivalence relation $\approx$ on A that satisfies, for all $a,b,c\in A$ and all $t\in Re$,

ii) $a\approx b \quad\leftrightarrow\quad a\odot c\approx b\odot c$,

iii) $a \simeq b \rightarrow t*a \simeq t*b$

(for short we write $a \simeq b$ instead of $(a,b) \epsilon \simeq$). It is generally accepted that, if $\simeq$ is interpreted as metameric color matching, properties ii) and iii) are empirically valid statements for quite a range of intensities and viewing conditions (but investigations of Zaidi and Pokorny (1973) indicate a failure of ii) for short-wave mixtures).

Let $X' = \langle A, \odot, *, \simeq \rangle$ and $Y' = \langle Re^m, +, \cdot, = \rangle$. Krantz (1975a) called the structure X' a Grassmann-structure and proved the existence of mappings ψ that map X' into Y' such that $a \simeq b$ iff $\psi(a) = \psi(b)$. The latter condition means that metameric lights are mapped onto the same vector over Re. He furthermore proved that, if trichromacy is accepted, the dimensionality of the vector-space Y' equals three and that the vector-homomorphisms ψ are unique up to positive nonsingular linear transformations. This means that $\psi = (\psi_1, \psi_2, \psi_3) \epsilon \Psi$ iff $(\psi'_1, \psi'_2, \psi'_3) = \psi' = \psi T \epsilon \Psi$ where T is a non-singular 3×3-matrix. Irtel (1982) generalized the representation and uniqueness theorem to binocular color matching.

The CIE recommends a specific representation $\psi \epsilon \Psi$ such that none of the coordinates ψ is negative and one of the coordinates represents luminance. This results, employing the projective transformation $\overline{\psi}_i = \psi_i / \sum_{j=1}^{3} \psi_j$, in the well-known xyz-chromaticity diagram.

There is, however, in such a representation no direct correspondence between the color matching functions ψ_i and the response characteristiques of the photoreceptors. If such correspondence has to be incorporated into the representation, further restrictions have to be imposed upon the ψ_i. In fact, this is equivalent to adding relations to X' such that the choice of T in $\psi' = \psi T$ is limited to strictly diagonal forms.

Vos and Walraven (1971; see also Walraven, 1974, and Vos, 1978a) succesfully used color matching data of the three types of color blinds to specify color matching functions that are unique up to a change of unit. Let $\asymp_i$ (i=1,2,3) denote color matching by a specific type of color blind. Then we know that $\asymp \subset \asymp_i$ and we know that the $\asymp_i$ also satisfy the above conditions ii) and iii). So the the structures $X'_i=\langle A,\odot,*,\asymp_i\rangle$ may be considered as reduced forms of X' and have two-dimensional linear vector spaces as their representations. On the basis of such orthogonal structures and information on wavelength discrimination and hue shift, Vos and Walraven were able to derive color matching functions proportional to response functions of the receptor systems. The results of colorimetry quite accurately describe equivalences among light stimuli. They do not, however, tell us which sensations of color are evoked by these stimuli. Quite prevailing in modern color theory is the notion that the four kinds of sensations of color - red, green, yellow and blue - arise as the result of two chromatic mechanisms: one signaling either "yellow" or "blue", the other signaling either "red" or "green". Theories of color perception that are based upon this idea are generally called "opponent colors theories" and what they aim at is the description of the actual sensations of color (or at least their presence or absence).

The significance of Krantz' (1975a,b) contributions to color science was that he not only created a measurement structure for color matching but enriched this structure with qualitative relations that were meant to incorporate data on the presence or absence of the sensations of red, green, yellow and blue. Enriching the measurement structure with additional relational constraints then amounts to introducing suitable relations or operations onto A such that they, on the one hand, correspond to perceptual relations and, on the other

hand, may be homomorphically mapped onto additional restrictions on X.

However, enriching a measurement structure by introducing new relations into the domain of discourse presents a problem which pertains to a kind of coherence of the relations involved and is generally called "meaningfulness".

Meaningfulness

In the context of measurement structures problems of meaningfulness may arise. It is, for eaxmple, well known that, if a mapping ψ of some X onto $<Re,\geq>$ exists, sentences like "$\psi(a)/\psi(b)=c\epsilon Re$" are not interesting since their correctness depends upon the particular choice of ψ. Because of the ambiguity of such sentences they are considered as meaningless.

This kind of ambiguity, and the related problem of appropriate statistics, inspired Stevens (1946) to distinguish between nominal, ordinal, interval and ratio scales and to consider statements in terms of ψ-values as meaningful if their content is invariant under the class of permissible scale transformations. Here the class of permissible scale transformations is a set of functions θ defined by $\theta\epsilon\Theta\leftrightarrow\{\psi\epsilon\Psi\rightarrow\theta\psi\epsilon\Psi\}$. This concept of meaningfulness may be easily formulated in the context of the theory of measurement structures: let $M=<X,Y,\Psi>$ be a measurement structure and S a k-ary relation on the numerical set of Y. Then S is meaningful iff, for all $\psi,\psi'\epsilon\Psi$

$$\{\psi(a_1),\ldots,\psi(a_k)\}\epsilon S \quad\leftrightarrow\quad \{\psi'(a_1),\ldots,\psi'(a_k)\}\epsilon S.$$

This idea about (quantitative) meaningfulness was extensively develloped by Adams, Fagot and Robinson (1965) and, probably, corresponds to more intuitive notions on meaningfulness that prevail among most researchers.

Two problems arise in connection with this idea. The first is that

meaningfulness, so defined, applies to numerical relations only, while those relations are only the result of and refer to thought and experiment on qualitative aspects of the domain of discourse. The second problem is that the representation itself may not be uniquely determined: $Y=\langle Re^{+},\geq,\cdot\rangle$ and $Y'=\langle Re,\geq,+\rangle$ are isomorphic, so any relational structure X for which Y is a representation, is also representable by Y'. In such a case the class of permissible scale transformations is not independent of the particular representation chosen.

Both of these problems would vanish if meaningfulness would be a concept that could (should) apply to relations of any relational structure and not just to those in numerical structures. Such a concept would then be qualitative in the sense that it would not necessarily refer to a numerical structure or a particular representation. Furthermore, such a concept would, because of its qualitative nature, be a criterion on the basis of which the relevance of data, as qualitative empirical relations, could be judged - only meaningful data to be considered relevant.

Below we will present a definition of meaningfulness that is qualitative in the above sense. It closely resembles definitions advocated by Pfanzagl (1968), Luce (1978) and Narens (1981). The definition amounts to a concept that applies to relations that are invariant under endomorphisms of their domain.

Definition 4. Let $Z=\langle C,T_1,\ldots,T_n\rangle$ be a relational structure and $\Gamma=\{\gamma,\gamma',\ldots\}$ be the set of endomorphisms of Z. Let T^* be a k-ary relation on C. Then T^* is meaningful in Z iff, for all $\gamma\in\Gamma$,

$$(c_1,\ldots,c_k)\in T^* \quad \leftrightarrow \quad \{\gamma(c_1),\ldots,\gamma(c_k)\}\in T^*$$

Note that in the above definition the relational structure Z is not restricted to be either numerical or non-numerical. This implies that meaningfulness may apply to an empirical relation - a data structure - as well as to a numerical relation.

The next argument shows that a non-primitive relation T^* that is meaningful in a given domain may be added to that domain such that its homomorphic image is meaningful in the representation and such that the original set of homomorphisms from the domain of discours into the representation is not altered.

Let $M=<X,Y,\Psi>$ be a measurement structure with $X=<A,R_1,\ldots,R_n>$ and $Y=<B,S_1,\ldots,S_n>$. Now suppose that R^* is a non-primitive k-ary relation on A that is meaningful in X and S^* a non-primitive k-ary relation on B that is defined by

$$(a_1,\ldots,a_k)\epsilon R^* \quad \leftrightarrow \quad \{\psi(a_1),\ldots,\psi(a_k)\}\epsilon S^*$$

for all $\psi\epsilon\Psi$. Next we define $M'=<X',Y',\Psi'>$ with $X'=<A,R_1,\ldots,R_n,R^*>$ and $Y'=<B,S_1,\ldots,S_n,S^*>$; i.e. the original relational structures enriched with relation R^* and S^*, respectively. Obviously, $\Psi'\subseteq\Psi$ and we also have, by the definitions of X', Y', S^* and Ψ, $\psi\epsilon\Psi\rightarrow\psi\epsilon\Psi'$ hence $\Psi=\Psi'$. Furthermore, for any fixed $\psi\epsilon\Psi$, if γ is an endomorphism of Y, $\gamma\psi\epsilon\Psi$ and so γ is an endomorphism of Y'. Obviously, each endomorphism of Y' is also an endomorphism of Y so the endomorphism-sets of Y and Y' coincide. Thus S is meaningful in Y.

It is interesting that the above argument also shows that endomorphisms of Y' are permissible scale transformations in the sense of Stevens (1946).

It should be understood that not every relation that is meaningful in the representation has a corresponding relation that is meaningful in the domain

of discourse. This would only be true if all the endomorphisms of the domain of discourse and of the representation are automorphisms (Luce, 1978; Theorem 2).

The definition given here is similar to one given by Narens (1981) and differs from that given by Luce (1978) in two respects. The first difference between the present and Luce's concepts of meaningfulness is that his approach does not recognize the fact that a relational structure may not contain a weak order as one of its primitives. This difference is not trivial since null-measurements in physics and matching procedures in psychophysics are abundant. The second difference concerns a generalization since Luce considers only automorphisms and all automorphisms are endomorphisms.

It is worth mentioning that, according to Definition 4, all the primitives of a relational structure are meaningful so one may choose any data-structure as one's domain of discourse, at the risk, however, of loosing contact with previously established measurement structures for more restricted domains of discourse.

A last remark on Definition 4 is intended to point to a somewhat paradoxical situation and demonstrates the arbitrariness of the level of abstraction chosen: an endomorphism of a relational structure X may be considered as a relation on the object set of X and one may wonder whether such a relation is meaningful in X. It is easy to see that if one requires it to be so, it should commute with every other endomorphism of $X=\langle A,R_1,\ldots,R_n\rangle$, i.e. if γ and γ' are endomorphisms of X they should satisfy $\gamma\{\gamma'(a)\}=\gamma'\{\gamma(a)\}$ for all $a\in A$. Apart from some representations in psychophysics (e.g. Chapter 2 of this manuscript) and psychometric theory (the Rasch-model, e.g. Van den Wollenberg, 1979), most representations employed in the social sciences only have non-commuting

endomorphisms (e.g. the linear transformation of interval scales).

The last two remarks stress the point that this concept of meaningfulness is arbitrary and, if accepted, only relevant with respect to a particular domain of discourse, the choice of which is arbitrary also, and with respect to a particular, but arbitrary, level of abstraction. We feel that this arbitrariness cannot be circumvented. At best, the heuristic value of this particular concept of meaningfulness may render this arbitrariness acceptable.

If new relations are introduced into a relational structure one should heed to the meaningfulness of such relations. The previously given definition obviously provides a tool to do so.

Meaningful data.

The main motive behind a position like Brindley's is the vague notion that states of the organism that are to be represented in physical or physiological units are somehow separated from states that are described in non-physical units.

Within the framework of measurement theory and the associated concept of meaningfulness Brindley's distinction between Class A and Class B observations seems rather obscure: data, as referring to relations and operations in an empirical domain, are either meaningful or not and their meaningfulness does not a priori depend upon the particular empirical domain but instead on the outcome of tests of axioms concerning the properties of these relations and operations.

Such an infertile separation between different domains of discourse and systems of units does not give heed to the existing and theoretically well-underpined (e.g. Palacios, 1964; Luce, 1978) system of units, the elements of

which indeed arise from different universes of discourse, that is generally accepted in other branches of science.

By now it should be clear that formulating a measurement theory and the possible enrichment thereof requires a concise and explicit notion of the body of phenomena under study. The testing of the axioms of the non-numerical structure, i.e. the object set of which the properties postulated are the primitive relations in X, then is equivalent to testing the theory. Once a measurement structure has been formulated the domain of discourse is clear and problems of linking different domains or "units" are, at least theoretically, solved for. The criterion of meaningfulness allows one to decide whether or not a measurement structure or an intended enrichment thereof constitute a meaningful class of observations.

Codes

Uttal (1967, 1973) was the first to introduce the concept of a neural code and Krantz (1975a) first used it in a formal sense. Uttal defined a code with respect to neural processes that can be discriminated behaviorally. Within the present context we prefer to undo it from the adjective "neural" since the connotations thereof may veil its general significance. The following definition is somewhat less direct than the one given in Chapter 2 but permits us to define completeness.

> *Definition 5.* Let $X=\langle A,E_0,\ldots,E_n\rangle$ be a relational structure such that, for each i, E_i is an equivalence relation on A, satisfying $E_0 \subseteq E_i$. Let $\{f_1,\ldots,f_n\}=F$ be a set of mappings from A into Re^+. F is called a set of codes iff, for all $a,b \in A$ and all i, $f_i(a)=f_i(b)$ iff aE_ib. F is called complete iff $E_1 \cap \ldots \cap E_n=E_0$.

This formalizes the earlier mentioned step from sensations to attributes of stimuli, the attributes being quantified by codes. For example, E_0 may be identified with complete indistinguishability, the other equivalences referring to specific sensations like brightness and hue.

It is worthwhile noting that, if $M=\langle X,Y,\Psi\rangle$ is a measurement structure, each homomorphism $\psi\in\Psi$ defines a code on the the object set of X.

Note further that the concept of a code, though general, is not entirely trivial: h_1 and h_2 may be measures without an associated equivalence relation while the code f_i may be constructed as $f_i=G(h_1,h_2)$ with $G:Re^2\to Re$.

The significance of codes may be realized by noting that every code is expressable as a function of a complete set of codes: let $F=\{f_1,\ldots,f_n\}$ be a complete set of codes and let g be a code such that $g\notin F$. Now define $H:Re^n\to Re$ for all $a\in A$ by

$$g(a) = H\{f_1(a),\ldots,f_n(a)\}.$$

Suppose $a,b\in A$ satisfy $f_i(a)=f_i(b)$ for all i. Then, since F is complete, aE_0b hence $g(a)=g(b)$. This proves H to be well defined. If we now accept the color-matching functions as a complete set of codes, a basis for relating the codes treated in the next chapters to the color-matching functions is formally provided.

Preview.

We are now in a position that allows to describe the content of the next chapters. The second chapter aims at describing a criterion to decide whether or not a specific paradigm - the assessment of the operating characteristics of an hypothetical detector through the determination of sensitivity functions

that describe response-equivalent operating environments - is meaningful relative to the structure <A,*>

In this chapter we present definitions of sensitivity functions that closely correspond to the actual procedures of construction as can be encountered in the literature on mechanisms of (color) vision As will appear, such constructions are unambiguous if, and only if, the equivalence relation on which they are based is constrained by a property called (matching-) "invariance"

Since endomorphisms τ of A in X=<A,*> are of the form $\tau(a)=t*a$ for any $t\in R^+$, equivalences E introduced into X should satisfy the condition aEb iff $t*aEt*b$ for all $t\in R^+$ and all $a,b\in$ A (invariance) If this condition is not met such an equivalence is, according to Definition 4, not meaningful in X Furthermore, this condition is necessary and sufficient to derive the existence of the sensitivity functions

Although the formal expression of these ideas within the framework of theories of measurement and meaning is new, their quintessence is, of course, not (e g Stiles, 1949, page 139) The present approach, however, offers a falsifiable hypothesis and illustrates, as is shown in chapters 3 and 4, that concepts that seem adequate to describe our sensations in everyday-language, like for example greenness, brightness and loudness, may not be fit to serve as unitary theoretical concepts in theories on sensations Chapter 2, apart from stating and proving a representation and uniqueness theorem for sensitivity functions in the structure <A,E,*>, offers a formal description of theories that relate certain kinds of codes to color-matching, the result being remarkably similar to the famous Pi-theorem from dimensional analysis (Buckingham, 1914, Thun, 1960, Luce and Narens, 1983) Indeed, these theories satisfy dimensional

invariance.

The last part of chapter 2 provides a formal description of null-measurement that abandons the requirement of additivity and formulates a sufficient condition on the kind of mutual dependence of the detectors involved that permits meaningful representation of the procedure.

Chapter 3 and 4 provide applications of the ideas formulated in chapter 2.

Chapter 3 concerns the question of meaningful measurement of the sensitivity for direct, heterochromatic matching of brightness. The first experiment establishes the meaningfulness of the matching relation and the remainder consists of the discussion of an application of the theory on the functional relation between color-matching and other codes. This discussion results in the tentative proposal of a model for direct brightness matching that behaves both super- and sub-additive in a manner that is, if not quantitatively, at least qualitatively in agreement with existing data.

The fourth chapter provides a test of invariance of yellow/blue cancellation, again exploiting the theoretical developments of the second chapter. It turns out that hue cancellation systematically deviates from invariance. It is then shown that earlier models of yellow/blue cancellation cannot satisfactorily describe all relevant data. A new account of these data is then provided on the basis of a model that predicts a new phenomenon concerning the wavelength shift of the short wave component of unique red. A test of this prediction does not disprove the new model.

Since yellow/blue cancellation does not satisfy invariance, it is not meaningful in the sense of Definition 4 and, consequently, the model describing the phenomenon is not dimensionally invariant. This casts serious doubt upon the fertility of approaches to hue perception that do not recognize

the ultimate untenability of "yellowness" and "blueness" as theoretical concepts.

The fact that concepts like "yellowness" and "blueness"are not fit to serve as theoretical concepts of course does not exclude the phenomena that give rise to them from scientific inquiry. If we want to enrich the color-matching space with relations that refer to sensations of hue we must devise empirical laws that describe these phenomena.

Among the most well-known examples of such an approach are the attempts of Jameson and Hurvich (1955) to describe the sensations of hue (H) and saturation (S) in terms of nonlinear transformations of color-matching functions (codes that are sensitivity functions for the cone systems) ψ_i:

$$H=\psi_1/(\psi_1+\psi_2) \quad \text{and} \quad S=\{|\psi_1|+|\psi_2|\}/\sum_i^3|\psi_i|$$

Here ψ_1 and ψ_2 are color-matching functions that are supposed to be identical to sensitivity functions of the opponent chromatic mechanisms.

Another example is the so-called CIE-UCS diagram that attempts at representing chromaticity-differences in a $(\bar{\psi}_i,\bar{\psi}_j)$- diagram (Wyszecki and Stiles, 1967). In this tradition the content of the fifth chapter should be placed. The paradigm is that of direct magnitude estimation and what is aimed at is a model that describes hue-veiling in terms of non-linear transforms of color-matching functions.

What more the next chapters have in common is that in each of them care has been taken to relate the phenomena studied to the operating characteristics of the cone systems. Of course it is a legitimate and often productive approach to construct theories on perceptual phenomena that are formulated entirely

without any reference to physical quantities but instead contain other (perceptual) variables as their primitives. An encompassing theory on visual phenomena, however, contains explicit links to the structure and functioning of the visual pathway. Such links then contain units that refer to dimensions that are meaningful in physical theory. The laws of physics and the dimensions in which they are stated satisfy the principle of dimensional invariance. Laws that relate physical and psychological variables in order to describe perceptual phenomena are, therefore, constrained to those that are also dimensionally invariant. So, Eisler's (1982) claim that any transformation of physical variables is permitted in order to formulate theories in psychophysics in fact implies a fundamental separation between theories on the "outer" and theories on the "inner" world.

Chapter 2

SENSITIVITY IN SPECTRAL STRUCTURES[1]

Abstract

Invariance of matching is shown to be necessary and sufficient for the existence of sensitivity functions that are unique up to multiplication by arbitrary positive real numbers Examples from auditory and visual psychophysics are discussed

Introduction

In many branches of science and engineering detectors play an important role, either because they are used as measurement devices or because the interaction between the detector and its operating environment is the object of scientific research These devices include detectors for radiant energy, sound or substances in solid or solute state and may be as different as a photodiode and the human eye, a microphone and the human ear or a mass-spectrometer and a taste-receptor

In each of these instances of a detector it is important to somehow characterize its operating capabilities Also, in each of these instances both the operating environment and the response of the detector can be described by specifying the amounts of physical quantities involved, like frequency of radiant energy or sound, concentration, exposure-time, temperature, current, conductance, etc

The behavior of the detector is then often described by one or more measures of its responsivity R=Y/X Here X stands for an amount of a quantity (partly) describing the operating environment and Y stands for an amount of

---

[1]C H Elzinga (1984) Sensitivity in spectral structures *Journal of Mathematical Psychology, 28,* 421-435

a quantity (partly) describing the response.

Silicon photodiode detectors, for example, produce an electric current in response to being exposed to radiant energy of appropriate power and wavelength. Their (absolute spectral) responsivity is then stated as an amount expressed in Amperes per Watt.

Thus one or more measures of responsivity characterize the operating capabilities of the detector.

If one is capable of unequivocally determining whether or not responses of a detector are equivalent the study of responsivity may be circumvented by specifying characteristics of the operating environment that evoke equivalent responses.

For example, a silicon photodiode detector may also be characterized by stating the power modulation that produces an amount of output quantity that is equivalent to the amount of output quantity that is caused by noise of fixed power and bandwidth: the so called N(oise) E(quivalent) P(ower) or its inverse, D(etectivity).

The measurement of responsivity requires that the dimensions of the quantities describing the operating environment and the response are known and firmly rooted in a theory on the behavior of the detector.

That such is not always the case is immediately clear if one is willing to consider a human observer as consisting of a set of detectors capable of discriminating different frequencies of sound or light with respect to loudness or brightness or of discriminating different substances with respect to smell and taste (note the lack of descriptive terminology for different qualities of smell like there is for different qualities of taste).

Meaningful measurement of responsivity requires response and operating environment measures that are unique up to scalar multiplication. For responses like the subjective sweetness of a solution, the loudness of a sound or the brightness of a light such measures do not at present exist.

Thus to specify the operating capabilities of the human observer responsivity measures have to be circumvented and, instead, equivalent response evoking characteristics of the operating environment must be specified.

That is in fact what has been done since the first decade of this century. Ives (1912a) for example, amply described the methods of flickerphotometry and direct brightness matching to determine equal brightness responses for differently coloured lights. Another example, also closely resembling the determination of NEP, is the determination of equal-loudness contours on a frequency axis (Fletcher and Munson, 1933). Such a contour provides the power of a sound of given frequency necessary to evoke a sensation of loudness, equivalent to that evoked by a certain amplitude of a fixed pure tone.

A limiting case of the use of equivalent responses is provided by the determination of threshold energy or threshold concentration. Such thresholds have been determined for most of the human senses under a variety of operating conditions.

In the sequel we shall use the term "(relative) sensitivity function" for functions on a set of operating environments that indicate to which degree the detector is affected and which are constructed through the use of equivalent responses. It will appear that the uniqueness of these sensitivity functions is difficult to state unless the detector behaves in a specific manner to be described below.

Spectral Structures and Codes

In each of the examples of a detector for radiant energy an important aspect of its operating performance is the way it does or does not discriminate between different frequencies or wavelengths. A frequency or wavelength interval is usually called a "spectrum" and operating characteristics that describe performance as a function of frequency are thus called spectral

responsivity or spectral sensitivity (functions). All spectra share the property that, at each point of the spectrum, the amplitude or power may be multiplied by a nearly arbitrary small or large number. Although frequency probably is not relevant for a taste receptor or a mass-spectrometer we still denote the structure that describes the operating environment as "spectral" since the amount of mass or concentration may also be manipulated by an operation that formally corresponds to scalar multiplication by positive real numbers.

Thus we characterize the possible operating environments of a detector as a set of different elements - different frequencies or substances - that is closed under scalar multiplication. The detector is then thought of as a device capable of discriminating between different elements of this set. These concepts are formally stated in

Definition 1. Let Re^+ denote the set of positive real numbers. The triple $\mathcal{A}=\langle A,E,*\rangle$ is a spectral structure if and only if A is a non-empty set that is closed under the scalar multiplication operation *; i.e. for all $a\in A$ and all $t,u\in Re^+$ we have

i) $1*a = a$

ii) $t*(u*a) = (tu)*a = u*(t*a)$

iii) $t*a\in A$

and E is an equivalence relation on A such that for all $a,b\in A$ there exists a unique $t\in Re^+$ that solves for $aE(t*b)$.

The expression aEb could represent observations like "the pitch of a matches to the pitch of b" or "the latency of the first spike in response to a matches to the latency of the first spike in response to b".

The statement that E is an equivalence relation severely limits the matching methodology. If reflexivity fails this may be due to, for example, differences in angle of incidence or area of contact. But then the detector has to be defined with respect to a specific area of contact or angle of

incidence. Transitivity might fail due to random background noise in which case some statistic could be more appropriate to define matching

The additional solvability condition imposed upon E is meant to ensure that the detector is sufficiently specific a pair of monochromatic lights or pure tones of sufficient frequency difference cannot be made indistinguishable to the human observer by manipulating the power of one member of the pair However, any such pair can be made to appear equally bright or equally loud by an appropriate change of the power of one of the members of the pair Lack of uniqueness of such power-settings could also be due to random noise and thus be circumvented by ensuring sufficient statistical precision

Of course, the statement that A is closed is an idealization since for very small or very large t detector activity will terminate The dynamic operating range for most human sense organs is, however, large enough to render this axiom (in Definition 1) acceptable

The detector, when affected by elements of A, produces responses. If one intends to characterize response-equivalent operating environments it is convenient to assign some numerical code or value to these elements. Following Krantz (1975a) we define a code as some real valued function on A that subserves discrimination with respect to the relevant dimension on which the detector operates More formally we state

Definition 2 Let $\mathcal{A}=\langle A,E,*\rangle$ be a spectral structure A function $f_{\mathcal{A}}$ that maps A onto Re^{+} is a code if and only if, for all $a,b\in A$

$$aEb \quad \leftrightarrow \quad f_{\mathcal{A}}(a)=f_{\mathcal{A}}(b). \qquad (1)$$

To construct a code for the spectral structure take a fixed, arbitrary element $a_o\in A$ and define $f_{\mathcal{A}}$ for all $a\in A$ by $aEf_{\mathcal{A}}(a)*a_o$ This is well defined, satisfies (1) and implies $f_{\mathcal{A}}(a_o)=1$ because E is reflexive and t in aEt*b is unique for all pairs $a,b\in A$

If and only if a code $f_{\mathcal{A}}$ satisfies $f_{\mathcal{A}}(t*a)=t^k f_{\mathcal{A}}(a)$ for some $k\in Re^+$ and all $a\in A$, $f_{\mathcal{A}}$ is called "homogeneous of degree k".

Sensitivity Functions in Spectral Structures

There are two experimental procedures that lead to functions on A that we shall call "sensitivity functions" although these functions are not always referred to by that name in the literature.

The first procedure amounts to obtaining the reciprocal of the power or concentration of elements of A required to elicit some fixed response from the detector. Some authors (e.g. Rodieck, 1973) call the resulting function (or its inverse) an "action spectrum", others (e.g. Brindley, 1960b) call it a sensitivity function and sometimes it is not given any special name (e.g. Grum and Becherer, 1979). To distinguish it from the function resulting from the second procedure we will call it a "constant response sensitivity function" and abbreviate this concept as "CRSF". Formally:

> *Definition 3.* Let $\mathcal{A}=\langle A,E,*\rangle$ be a spectral structure and $f_{\mathcal{A}}$ a code on A. The function S^m from A into Re^+ is a CRSF if and only if, for some $m\in Re^+$, all $a\in A$ and all $t\in Re^+$,
>
> $$S^m(a) = t^{-1} \quad \leftrightarrow \quad f_{\mathcal{A}}(t*a) = m$$

Note that each S^m is a code on A too and that, because of its definition and the associativity of *, $S^m(t*a)=tS^m(a)$ for all $a\in A$ and all $t\in Re^+$, i.e. it is homogeneous of first degree.

The second procedure amounts to selecting some arbitrary element c from A as a standard or unit element and comparing all other elements of A with that standard by determining a $t\in Re^+$ such that $aE(t*c)$. We will call the resulting function a "fixed standard sensitivity function", abbreviated as "FSSF". Formally:

> *Definition 4.* Let $\mathcal{A}=\langle A,E,*\rangle$ be a spectral structure and $f_{\mathcal{A}}$ a code on

A. The function S^c from A into Re^+ is a FSSF if and only if, for all $a \epsilon A$, all $t \epsilon Re^+$ and some $c \epsilon A$,

$$S^c(a) = t \quad \leftrightarrow \quad f_{\mathcal{A}}(a) = f_{\mathcal{A}}(t*c)$$

Note that S^c is a code also and that it is not, contrary to S^m, necessarily homogeneous.

The dual nature of CRSF and FSSF is immediately clear when one realizes that, although these concepts are defined through the use of codes, they could have been defined without any reference to the concept of a code: simply define $S^c(a)=t$ iff $aE(t*c)$ and $S^m(a)=t^{-1}$ iff $(t*a)Ec$ and the symmetry of the two procedures is clear.

The procedure described in Definition 3 is used much more frequently than the procedure leading to a FSSF, probably because it seems to be specially suited for the measurement of thresholds of detection and determination of thresholds is generally more reliable than supra-threshold matching.

Stevens and Davis (1938) describe the FSSF-procedure for pure tones with c a pure tone of 1000 Hz and call the resulting values "loudness-levels". Jameson and Hurvich (1955) used the FSSF-procedure for measuring sensations of redness, greenness, blueness and yellowness of narrow-band wavelength distributions and called the resulting functions "cancellation functions".

For many purposes it is convenient to transform a sensitivity function to a dimensionless quantity that is called a "relative sensitivity function" (for short: RSF). For the sake of simplicity of argument and notation we will first restrict our attention to the FSSF and its associated RSF:

Definition 5. Let $\mathcal{A}=<A,E,*>$ be a spectral structure and S^c be a FSSF for $\mathcal{A}$. Then the function U_b^c is an RSF for $\mathcal{A}$ if, for some $b \epsilon A$ and all $a \epsilon A$,

$$U_b^c(a) = S^c(a)/S^c(b)$$

The spectral structure is called "proper" if, for all $d \epsilon A$, $U_b^c = U_b^d$

Properness, though a property of the spectral structure, will also be used as a term referring to the particular RSF's associated with a proper spectral structure. This should not lead to ambiguity.

The element b from Definition 5 is often referred to as the "normalization element", $S^c(b)$ as the "normalization factor" and the procedure of constructing an RSF as "normalizing". Essentially, the choice of the normalization element is arbitrary although, depending upon the empirical domain, some choices may be more appropriate then others.

Properness of the spectral structure is important since each U_b^c induces a weak order D_b^c on A: for all $a,d \epsilon A$ define the asymmetric part of the weak order as $aO_b^c d$ iff $U_b^c(a) > U_b^c(d)$ and the symmetric part by $aI_b^c d$ iff $U_b^c(a) = U_b^c(d)$ and $D_b^c = O_b^c \cup I_b^c$. If one RSF is proper, all of the so defined weak orderings are identical and it is then meaningful to interpret an RSF as a measure of detector activity. In the next paragraph we will, therefore, present a theorem that precisely states a property of the code $f_{\mathcal{A}}$ that is necessary and sufficient for the RSF's associated with $\mathcal{A}$ to be proper.

Properness and Invariance

> *Theorem 1.* Let $\mathcal{A} = \langle A, E, * \rangle$ be a spectral structure, $f_{\mathcal{A}}$ a code on A, F a function from A onto Re^+ such that for all $a \epsilon A$ and all $t \epsilon Re^+$ $F(t*a) = tF(a)$ and L is a one-to-one function from Re^+ onto itself. Then a detector generates a proper RSF if and only if its associated code satisfies, for all $a \epsilon A$,
>
> $$f_{\mathcal{A}}(a) = L[F(a)] \qquad (2)$$

Proof of Theorem 1. (Necessity) Suppose $f_{\mathcal{A}} = L(F)$ where L and F are defined as in Theorem 1. Then, because of Definition 4 and the definition of F we have

$$S^c(a) = t \quad \leftrightarrow \quad L\{F(t*c)\} = L\{F(a)\}$$

$$\leftrightarrow \quad F(t*c) = F(a) = tF(c)$$

$$\leftrightarrow \quad t = F(a)/F(c)$$

for all $a\epsilon A$ and all $t\epsilon Re^+$. Now we easily obtain through Definition 5, for $c\neq d$

$$U_b^c(a) = \frac{F(a)/F(c)}{F(b)/F(c)}$$

$$= \frac{F(a)/F(d)}{F(b)/F(d)} = U_b^d(a)$$

(Sufficiency) We need to show that L and F can be constructed according to their definition in the Theorem and that $f_{\mathcal{A}}=L(F)$ if U_b^c is proper.
Assume the spectral structure is proper, i.e. there exist $b,c\epsilon A$ such that $U_b^c=U_b^b$ hence

$$S^c(a)/S^c(b) = S^b(a) \quad \leftrightarrow \quad S^c(a) = S^c(b)S^b(a) \qquad (3)$$

since $S^b(b)=1$. Because of (3) we have[2]

$$S^c(t*a) = S^c(a)S^a(t*a) = tS^c(a) \qquad (4)$$

for all $a\epsilon A$ and all $t\epsilon Re^+$. Then the construction of F is accomplished by using a given, fixed $c\epsilon A$ and defining $F(a)=S^c(a)$ for some fixed c. Now the existence of L such that $f_{\mathcal{A}}(a)=L\{F(a)\}$ for all $a\epsilon A$ is guaranteed since F completely determines $f_{\mathcal{A}}$:

$$F(a) = F(b) \quad \leftrightarrow \quad S^c(a) = t = S^c(b)$$

$$\leftrightarrow \quad f_{\mathcal{A}}(a) = f_{\mathcal{A}}(t*c) = f_{\mathcal{A}}(b).$$

---

[2]Kruithof (1953) already suggested to test the right hand side of (3) for direct brightness matching in order to establish whether brightness is to be considered as an "...univalent function of chromaticity...".

This completes the proof. □

Actual construction of L could be accomplished by taking $L(t)=f_{\mathcal{A}}(t*c)$ for all $t\in Re^{+}$ and fixed c. Then $L(t)=L\{S^{c}(t*c)\}=L\{F(t*c)\}$ for every $a\in A$ with $F(a)=t=F(t*c)$.

Before we treat RSF's based upon CRSF's we will elaborate upon the proof of Theorem 1.

In proving the sufficiency of properness for the existence of F and L as demanded by the Theorem, we obtained (4); i.e. homogeneity (of first degree) of the sensitivity function. Except for the derivation of (4) we did not use properness, so homogeneity of S^{c} is a sufficient condition for the properness of the RSF. Obviously, homogeneity is also necessary for properness since it implies u=v if $S^{c}(a)=uS^{d}(a)$ and $S^{c}(b)=vS^{d}(b)$ for any $a,b\in A$.

Note further that homogeneity of S^{c} implies for all $a,b\in A$ and all $t\in Re^{+}$,

$$f_{\mathcal{A}}(a) = f_{\mathcal{A}}(b) \quad \leftrightarrow \quad f_{\mathcal{A}}(t*a) = f_{\mathcal{A}}(t*b) \qquad (5)$$

since

$$f_{\mathcal{A}}(a) = f_{\mathcal{A}}(b) \quad \leftrightarrow \quad S^{c}(a) = S^{c}(b)$$

$$\leftrightarrow \quad S^{c}(t*a) = S^{c}(t*b)$$

$$\leftrightarrow \quad f_{\mathcal{A}}(t*a) = f_{\mathcal{A}}(t*b).$$

Codes that satisfy (5) are called "invariant" (Krantz, 1975a). It is easily demonstrated that invariance of $f_{\mathcal{A}}$ implies first degree homogeneity of S^{c} so invariance of the code itself is a necessary and sufficient condition for properness of the RSF.

Note that the "multiplicative" structure of the code is implicit in (2): from (2) we obtain $F(a)=L^{-1}\{f_{\mathcal{A}}(a)\}$ so $f_{\mathcal{A}}(t*a)=L[tL^{-1}\{f_{\mathcal{A}}(a)\}]$.

At first sight one might think that first degree homogeneity of S^{c} is too

strong a requirement since if $S^c(t*a)=g(t)S^c(a)$, where g is some function, would also lead to properness. However, such a relaxation would lead to $t=S^c(t*c)=g(t)S^c(c)=g(t)$ so either $g(t)=t$ for all $t\in Re^+$ or the uniqueness of t in $aE(t*c)$, as required by Definition 1, would be violated.

We could introduce a formal addition on A:

$$f_{\mathcal{A}}(a) = f_{\mathcal{A}}(t*c),\ f_{\mathcal{A}}(b) = f_{\mathcal{A}}(u*c) \quad \rightarrow \quad a\odot b = (t+u)*c$$

for all $a,b\in A$ and all $t,u\in Re^+$. $\odot$ is well defined, associative, commutative and for all $t\in Re^+$ we have $t*(a\odot b)=(t*a)\odot(t*b)$. It is easily seen that, with $\odot$ as defined above $S^c(a\odot b)=S^c(a)+S^c(b)$. Since it is not guaranteed that aEb iff $(a\odot d)E(b\odot d)$ for all $a,b,d\in A$, addition of sensitivities might not have any clear empirical interpretation.

Obviously, the code $f_{\mathcal{A}}$ is invariant if and only if

$$aEb \quad \leftrightarrow \quad (t*a)E(t*b), \qquad (6)$$

$a,b\in A$ and all $t\in Re^+$. In the sequel we will refer to any symmetric binary relation that satisfies (6) as "invariant".
Equally obvious, since homogeneity of S^m is independent of invariance of E this invariance is necessary and sufficient for the identity of S^c and S^m.

Analoguous to Definition 5 we now define an RSF based upon the concept of a CRSF:

Definition 5'. Let $\mathcal{A}=\langle A,E,*\rangle$ be a spectral structure and S^m be a CRSF for $\mathcal{A}$. Then the function U^m_b from A into Re^+ is an RSF if, for some $b\in A$ and all $a\in A$,

$$U^m_b(a) = S^m(a)/S^m(b).$$

U^m_b is called "proper" if for all $n,m\in Re^+$ $U^m_b=U^n_b$.

Of course, U^m_b is proper if and only if E is invariant. So, since the

construction of a code for a spectral structure is always possible we have proven the following

> *Corollary.* Let $\mathcal{A}=\langle A,E,*\rangle$ be a spectral structure and U, U' be a pair of RSF's for $\mathcal{A}$. Then there exists a constant $v\in Re^{+}$ such that $U'(a)=vU(a)$ for all $a\in A$ if and only if E satifies invariance.

It is not difficult to see that a proper spectral structure satisfies conditions such that automorphism-meaningfulness, endomorphism-meaningfulness and N-quantitative meaningfulness coincide (Narens, 1981, Theorem 3.2). This justifies the previous employment of the term meaningful in connection with proper RSF's.

Note that the previous results sharply contrast with Brindley's assertion that "...invariance, however, is no essential part of the concept of a spectral sensitivity." (Brindley, 1960b, page 110).

Invariance of E is a qualitative, testable property of E and we will discuss applications of such tests and some of the results of the next paragraph on compound detectors and their sensitivity functions.

Compound Detectors

In several areas detectors are studied of which the response may be thought of as depending upon the responses of a finite set of more basic or primal detectors. We will call such detectors "compound".

For example, judgments about equality or inequality of brightness of different patches of light are thought to be mediated by the responses of three cone systems, and possibly the rod system, of the human eye. The RSF's of these systems are (approximately) known and the problem may then arise of how to describe brightness sensitivity in terms of these RSF's (e.g. Guth, Massof and Benzschawel, 1980). Similar problems have arisen in the study of hearing where responses from two ears combine (e.g. Levelt, Riemersma and Bunt, 1972). Let U be a RSF for a compound detector and

$u(a)=\{u_1(a),\ldots,u_n(a)\}$ a finite set of proper RSF's on which U depends; i.e. for all a,bϵA we have

$$U(a) = F\{u(a)\} \qquad (7)$$

where F is an arbitrary function. If the u_i are supposed to be proper, U is homogeneous of first degree if and only if F is homogeneous of first degree; i.e. $F\{tu(a)\}=tF(u(a)\}$ for all for all aϵA and all t$\epsilon$$Re^+$.

Sirovich (1977) specified F in (7) by assuming that U is proper and $\lim_{t\to 0}\{U^v(t*a)\}$ is differentiable with respect to t for all aϵA and some rational number v and therefrom derived

$$U(a) = \{\sum_{i=1}^{n} w_i u_i^v(a)\}^{1/v} \qquad (8)$$

where the w_i are arbitrary non-negative constants.

It is, however, easy to see that, assuming F homogeneous of degree 1 and the u_i proper,

$$U = F\{u\} = u_j F(u_1/u_j,\ldots,1,\ldots,u_n/u_j)$$

$$= u_j G(u_1/u_j,\ldots,u_n/u_j) \qquad (9)$$

Since, in the above equation, G is an arbitrary function, we may rewrite (9) as

$$U = \{\sum_{i=1}^{n} w_i u_i^v\}^{1/v} H(u_1/u_j,\ldots,u_n/u_j) \qquad (10)$$

So the differentiability assumption of Sirovich implies that H is constant. Although the validity of the differentiability assumption Sirovich imposed upon U is hard to test empirically, a consequence of (8) can be tested qualitatively. In order to explain this we introduce a rather natural operation $\odot$ on A such that $\langle A,\odot\rangle$ is a commutative semigroup; a$\odot$b may then be interpreted as a wavelength-by-wavelength energy sum or a sum of masses or concentrations.

Now suppose that the u_i are all linear, i.e. $u_i(t*a\odot s*b)=tu_i(a)+su_i(b)$ for

all a,b$\in$A and all s,t$\in$$Re^{+}$. Then (8) implies that, for v>1, U is sub-additive; i.e. $U(a \odot b) \leq U(a)+U(b)$, and, if v<1, U must be super-additive: $U(a \odot b) \geq U(a)+U(b)$ for all a,b$\in$A. Thus, if U is proper and the u_i are proper and linear, super-additivity for some a$\odot$b and sub-additivity for some other c$\odot$d, implies that the differentiability assumption must be rejected.

Some applications.

The earliest psychophysical experiment I am aware of that may be considered as a test of invariance of perceptual matching was performed by Lemberger (1908). She determined equally sweet concentrations for sucrose and crystallose and found the concentration ratio gradually increasing with increasing concentration of sucrose. Cameron (1947) demonstrated the same phenomenon for various other sweet substances which indicates that no proper RSF for sweetness exists and thus the ordering induced by an RSF for sweetness is not unique. This is reflected in the fact that power functions for magnitude estimation of sweetness have exponents that depend upon the substances involved (Frijters and Oude Ophuis, 1983).

Fletcher and Munson (1933) and Robinson and Dadson (1956) determined spectral sensitivity functions for loudness at different levels of sound pressure. If invariance of loudness matching were to hold, these functions ought to be parallel; they are definitely not, especially not so in the low-frequency range. Thus, loudness matching also violates invariance.

Cavonius and Hiltz (1973) found that magnitude estimates for brightness of monochromatic lights can be described by powerfunctions with a constant exponent for all wavelengths. This indicates that direct brightness matching could be invariant. Elzinga and De Weert (1984a)[3] explicitly tested for invariance of direct brightness matching. They presented their subjects with

---

[3]Reproduced as Chapter 3 of this manuscript.

two differently coloured lights in a bipartite field with the radiance of one half of the field fixed. The subjects had to adjust the radiance of the other half of the field such that both halves were perceived as "equally bright". Thus if a and b denote lights of different wavelength composition and s and t denote the radiances of a and b respectively, the subject determined t in s*aEt*b where E denotes equality of perceived brightness. Obviously, t depends upon a, b and s and E is invariant if and only if $t=f_a^b(s)$ is of the form $k_a^b s=f_a^b(s)$ where k_a^b is a constant depending upon a and b alone. Figure 1 shows data of a typical subject in log-log coordinates.

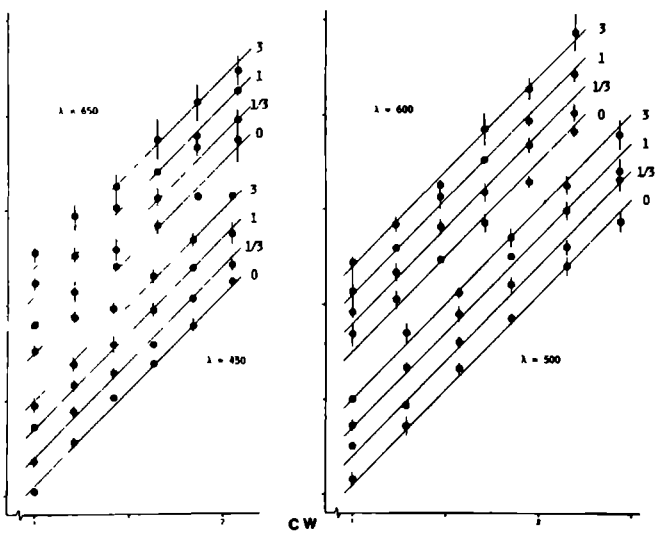


Figure 1. Equal brightness data of one subject from Elzinga and De Weert (1984a). The abcissa represents log-radiance of a reference light of 550 nm; the ordinate represents log-radiance of bichromatic mixtures of 550 nm and four different addend wavelengths in four different mixture-ratios. The addend wavelenghts and mixture ratio's are indicated in the figure. The points represent the means of 5 replications, the vertical bars the 80%-confidence intervals for those means. The drawn lines represent the best fitting linear regression lines. The regression lines and the corresponding points have been arbitrarily displaced along the ordinate.

Evidently, the data confirm invariance of direct brightness matching and thus RSF's for direct brightness matching are unique up to multiplication by an arbitrary positive constant.

Since human photopic vision depends upon stimulation of three independent cone systems, the human brightness detector may be thought of as a compound detector. The RSF's of the three cone systems are generally thought to be linear so an RSF for direct brightness matching can be written in the form of (10). Since direct brightness matching is both super- and sub-additive the function H appearing in (10) cannot be constant.

Let U denote an RSF for direct brightness matching and u_1, u_2, u_3 the RSF's for the cone systems with maximal sensitivities in the short, middle and long wavelength region respectively. Elzinga and De Weert (1984a) proposed (see also Chapter 3, formula 7)

$$U = \sum_{i=1}^{3} w_i u_i \{x_1 u_2/u_1 + x_2 u_3/u_2\}^v \qquad (11)$$

where the w_i and x_i are positive constants and $v \in (0,-1)$. This model is consistent with (10), nicely fits RSF's for direct brightness matching and data on both super- and sub-additivity. Furthermore, the model has an interpretation that is in qualitative agreement with current theories on brightness perception (Boynton, 1979): the linear part of (11) may be conceived of as representing an achromatic detector and the nonlinear part as representing the contribution of two chromatic detectors.

Cancelling Detectors

Most elementary textbooks on chemistry describe and explain the principle of titration: the measurement of acidity through determining the amount of alkaloid that has to be added in order that the resulting solution is neutral.

The quality of acidity (free hydrogen ions) is thus cancelled by the

different quality of the alkaloid (free hydroxide ions) and the resulting solution has neither of these qualities. We may conceive of the set of fluids A as being composed of three mutually disjoint subsets A_1, A_2 and A_3: the acids, the alkaloids and the "neutrals".

In color vision a similar kind of cancellation occurs: the quality of greenness of a light may be cancelled by admixing an appropriate amount of a reddish light and the resulting mixture does have neither of these properties (Jameson and Hurvich, 1955; Larimer, Krantz and Cicerone, 1974; Elzinga and De Weert, 1984b). Here, the response of a "redness detector" is thought to be cancelled by the response of a "greenness detector" and as a measure of the activity of one detector the amount of stimulation of the other detector is used.

The concept of a physical structure that allows for cancellation may be formally stated as

Definition 6. Let $\langle A,\odot\rangle$ be a commutative semigroup and * denote scalar multiplication on $\langle A,\odot\rangle$. Furthermore, let A have non-empty subsets A_1, A_2 and A_3 such that $A=A_1 \cup A_2 \cup A_3$ and $A_i \cap A_j=\emptyset$ for $i\neq j$ and let $C \subset A_1 \times A_2$ be a symmetric binary relation such that, for i,j=1,2 and $i\neq j$

$$aCb \quad \leftrightarrow \quad a\in A_i,\ b\in A_j,\ a\odot b\in A_3$$

Then the quadruple $\mathcal{A}=\langle A,C,\odot,*\rangle$ is a cancellation structure if and only if, for all $a\in A_i$ and all $b\in A_j$ there exists a unique $t\in Re^+$ such that $aC(t*b)$.

Cancellation may then be quantified by introducing the concepts of a cancellation function (for short: CF) and a relative cancellation function (for short: RCF):

Definition 7. Let $\mathcal{A}=\langle A,C,\odot,*\rangle$ be a cancellation structure. Then the function C^a from A_i (i=1,2) onto Re^+ is a CF if, for all $a\in A_j$, all $b\in A_i$ ($i\neq j$; i,j=1,2) and all $t\in Re^+$

$$C^a(b) = t \quad \leftrightarrow \quad bCt*a$$

Definition 8. Let $\mathcal{A}=<A,C,\odot,*>$ be a cancellation structure and C^a a CF on A_i and let $c\epsilon A_i$ (i=1,2). The function W_c^a is an RCF if, for all $b\epsilon A_i$

$$W_c^a(b) = C^a(b)/C^a(c)$$

The cancellation structure is called "proper" if and only if, for all $a,a'\epsilon A_j$, $W_c^a=W_c^{a'}$.

Again we eventually refer to both cancellation structures and their associated RCF's as being proper.

Now suppose $\mathcal{A}_i=<A_i,E_i,*>$ is a spectral structure. Just like an RSF, each RCF on A_i induces a weak order on A_i so properness of an RCF is sufficient for the identity of these weak orders. Furthermore, if the RCF is also to be considered as a measure of detector activity the unique order induced by a proper RCF on A_i should be identical with the order induced by a proper RSF. The next theorem provides an axiom that is sufficient for the orderings induced by proper RSF's and proper RCF's to be identical.

Theorem 2. Let $\mathcal{A}=<A,C,\odot,*>$ be a cancellation structure such that $\mathcal{A}_i=<A_i,E_i,*>$, i=1,2, are spectral structures. If, for all $a\epsilon A_j$ and all $b,c\epsilon A_i$ (i,j=1,2; $i\neq j$)

iv) $aCb \rightarrow \{ aCc \leftrightarrow bE_ic \}$

then the E_i (i=1,2) are invariant if and only if C is invariant.

Proof of Theorem 2. (a) Suppose C is invariant and, for some $b,c\epsilon A_i$ we have bE_ic. Now for fixed $a\epsilon A_j$ there exists a unique $t\epsilon Re^+$ such that $bC(t*a)$. Then, because of iv) we have $cC(t*a)$. Because C is invariant we now have $(u*b)C(ut*a)$ and $(u*c)C(ut*a)$ for all $u\epsilon Re^+$ and, using iv) again, we also have $(u*b)E_i(u*c)$ for all $u\epsilon Re^+$ so E_i is invariant.

(b) Suppose E_i is invariant, $a,b\epsilon A_i$ and $c,d\epsilon A_j$. Now there exist unique

$w,v,x,y \in Re^+$ such that $bC(w*c)$, $bC(v*d)$ and $aC(y*c)$, $aC(x*d)$. Using iv) and invariance of E_i we must have $\{(w/v)*c\}E_i d$ and $\{(y/x)*c\}E_i d$ so $w/v=y/x$. Let C^c and C^d be CF's for A_i; then we have

$$C^c(b)/C^d(b) = C^c(a)/C^d(a) \qquad (12)$$

Because of (12) we obtain, for all $a \in A_i$ and all $t \in Re^+$,

$$C^c(t*a) = \{C^d(t*a)/C^d(a)\}C^c(a).$$

For each pair (a,d) with $a \in A_i$ and $d \in A_j$ the ratio $C^d(t*a)/C^d(a)$ depends upon t only, hence

$$C^c(t*a) = g(t)C^c(a).$$

Because of the associativity of *, g must satisfy $g(uv)=g(u)g(v)$ for all $u,v \in Re^+$. This is an ordinary Cauchy-equation with solution $g(t)=t^k$ for all $t \in Re^+$ and arbitrary real k (Aczél, 1966). Consequently, aCc implies $(t*a)C(t^k*c)$ for all $a \in A_i$, all $c \in A_j$ and all $t \in Re^+$. But C is symmetric hence $k=1$. This completes the proof. □

Note first that (12) implies $W_b^c=W_b^d$ for all $c,d \in A_j$ and all $b \in A_i$ so under the conditions of Theorem 2 both invariance of E_i and invariance of C imply properness of the RCF's on A_i.
Furthermore, under the conditions of Theorem 2, invariance of either C or E_i implies that for all $b \in A_j$, $U_c^a=W_c^a$ for each $c \in A_i$ hence the orderings on A_i induced by RSF's and RCF's are identical and thus refer to the same quality of the elements of A_i.

Iso-cancellation of yellowness/blueness.

Let A_1 denote the set of all yellowish lights, A_2 the set of all bluish lights and A_3 the set of all lights that are not perceived as either yellowish or bluish. Cancellation of yellowness of a light patch is always possible by admixing an appropriate amount of bluish light and vice versa.

This is known since long and has become part of the qualitative basis of many models in color vision. The phenomenon was first systematically observed and accounted for by Jameson and Hurvich (1955). Later on, Larimer, Krantz and Cicerone (1974) explicitly tested for the linearity of the yellow/blue and red/green cancellation functions and found weak evidence for violations of linearity for the yellow/blue

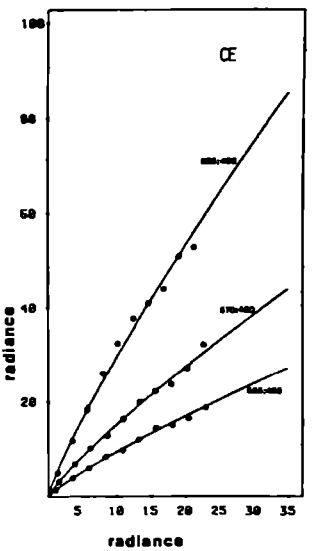


Figure 2. Iso-cancellation curves for yellowness/blueness for three different mixtures of yellowish and bluish monochromatic lights (Elzinga and De Weert, 1984b). The wavelenghts of the components of each mixture are indicated in the Figure. The horizontal and vertical axis represent the radiance of the longer, yellowish wavelength and the shorter, bluish wavelength respectively, both in the same, arbitrary unit. The dots represent the means of 8 replications, the drawn lines were generated from the assumption that the iso-cancellation functions are power functions.

cancellation functions. Elzinga and De Weert (1984b)[*] systematically tested for invariance of yellow/blue cancellation through determining what they

---

[*]Reproduced as Chapter 4 of this manuscript.

called the "iso-cancellation function": a mapping h_b^a from Re^+ into Re^+ defined by $s*aCh_b^a(s)*b$. Again, if and only if $h_b^a(s)=k_b^a s$ where k_b^a is a constant depending on $a\epsilon A_i$ and $b\epsilon A_j$ only, C is invariant. Figure 2 shows the data of one typical subject for three different mixtures of yellowish and bluish monochromatic lights. The subjects in this experiment were presented with a mixture $s*a\oplus t*b$ of fixed radiance s and variable radiance t and were instructed to manipulate the radiance t such that the mixture was perceived as neither yellowish nor bluish; i.e. such that $t=h_b^a(s)$.

From Figure 2 it is evident that invariance of C is violated in all three wavelength combinations. This indicates that proper RSF's for yellowness and blueness, if at all existent, cannot be determined through cancellation.

Chapter 3.

INVARIANCE OF DIRECT BRIGHTNESS MATCHING[1]

Abstract

It is demonstrated that brightness sensitivity functions are self-similar, i e unique up to multiplication by a positive constant. This invariance leads to a structural restriction for sensitivity models that are formulated in terms of action spectra of cone systems. The heuristic value of this restriction is demonstrated through simulations on various kinds of data on direct brightness matching and brightness magnitude estimation on the basis of a new model for brightness sensitivity that both accounts for super- and sub-additivity.

Introduction

By now the validity of the concept of brightness has been widely accepted: spectral luminosity functions determined by direct brightness matching are different from those obtained by flickerphotometry (Ives, 1912a; Dresler, 1953; Wagner and Boynton, 1972), brightness matching is non-additive (Ives, 1912c; Kohlrausch, 1935; MacAdam, 1950) as opposed to flickerphotometric matching (Ives, 1912c; Ikeda, 1983) and the magnitude of the brightness sensation seems to be a power function with an exponent that is wavelength-independent (Ekman, Eisler and Künnapas, 1960; Cavonius and Hiltz, 1973; Mansfield, 1973, 1976). Furthermore, spectral sensitivity data for the brightness mechanism have been published by several authors (Ives, 1912a,b,c; Wilson, 1964, Comerford and Kaiser, 1975).

The present investigation was undertaken in order to establish whether or

---

[1]Submitted for publication in the Journal of the Optical Society of America

not a proper relative sensitivity function (for short: RSF) can be unambiguously attached, i.e. without reference to an arbitrary standard response or standard stimulus, to the concept of brightness.
In order to explain the logic of the first experiment we will elaborate on the concept and construction of an RSF in a slightly formal manner.

Spectral Sensitivity and Invariance

For the construction of sensitivity functions two different procedures are commonly used. The first procedure amounts to selecting an arbitrary standard element from the domain of interest and comparing all other elements with that standard through determining the radiance of the standard required to match those elements; that radiance is then considered to be the sensitivity at the element concerned.
The second procedure consists of selecting an arbitrary response level and determining sensitivity as the reciprocal of the radiance required of any domain element to evoke that level of response.
It is, furthermore, quite common to normalize the so obtained functions to some arbitrary value at a maximal or minimal point. Such normalized functions are then called RSF's.
However, neither of the procedures in and by itself guarantees that the resulting functions have an unambiguous interpretation, since it is by no means certain that the selection of a different standard element or a different response level would lead to the same RSF, let alone that the results of the two different procedures are identical up to multiplication by a positive constant (self-similar).

Elzinga (1984)[2] proved that the set of all possible RSF's that could be obtained by either of the two procedures are self-similar if and only if the matching-behavior satisfies a property called 'invariance' (Krantz, 1975a). More formally, let A denote the domain with elements a,b∈A and let t* denote the operation of changing the radiance level with some positive constant t. Furthermore, let aEb denote the observation that (the response evoked by) a matches to (the response evoked by) b. Then the matching-behavior satisfies invariance if, for all positive t and all a,b∈A

$$aEb \quad \leftrightarrow \quad (t*a)E(t*b) \qquad (1)$$

If the matching-behavior does not satisfy invariance the concept of sensitivity does not apply since the measurement procedure leads to arbitrary results. This is, for example, the case for the measurement of luminance-sensitivity through flickerphotometry (Dresler, 1953; Ingling, Tsou, Gast, Burns, Emerick and Riesenberg, 1978; Yaguchi and Ikeda, 1983). This proposition sharply contrasts with Brindley's (1960a) opinion that '... invariance, however, is no essential part of the concept of a spectral sensitivity' and it stresses the fact that the employment of physical units by itself does not guarantee the meaningfulness of the results of a measurement procedure.

Experiment I: Invariance of brightness

The literature contains several indications that (1) might hold indeed for

---

[2]Reproduced in Chapter 2.

brightness matching.

Ives (1912a) determined RSF's for brightness at different radiance levels of the standard stimulus and the resulting curves seem to have approximately the same shape. He (Ives, 1912c) also determined RSF's with differently coloured standards and these RSF's appear to be self-similar.

Although Ives presented his data rather extensively, his publications do not contain the necessary details that would allow for a thorough test of (1).

A second indication that invariance might hold for direct brightness matching can be found in the results of Wilson (1964). She obtained an equal-detectability curve that has approximately the same shape as the RSF for brightness. This result, reproduced by Estévez (personal communication), was used by Krantz (1974) as an argument in favour of (1).

A third indication that brightness might be an invariant code is provided by the fact that metameric matches are generally believed to satisfy invariance: by definition, metamers are equally bright. This indication is weak since direct brightness matching requires the subject to abstract from irrelevant aspects like differences in saturation and/or hue whereas metameric matching is essentially a perturbation experiment.

The experiment presented here intends to be a direct and detailed test of (1), both for monochromatic lights and bichromatic mixtures.

Let ⊙ denote the operation of additive color mixture. The first experiment then consists of determining t for different values of s and q in

$$(s*a)E[t*\{q*a\odot(1-q)*b\}], \quad 0 < q < 1 \qquad (2)$$

where a and b are monochromatic. If, and only if (1) holds for all a, b, q the functions $f(s|a,b,q)=t$ must be linear in the plane (s,t) with the

intercept at the point (0,0).

Apparatus

A schematic diagram of the optical equipment is shown in Figure 1. Three halogen tungsten ribbon filament lamps (220V, 150W, underrun at 0.6A, AC-stabilized voltage) S1, S2 and S3 were used as light sources. Light from each of these sources entered Oriel grating monochromators (M1, M2 and M3), providing the monochromatic beams.

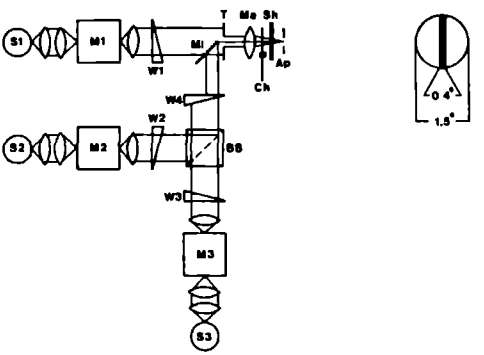


Figure 1. Diagram of the apparatus. S1, S2 and S3 are 150 W halogen tungsten ribbon filament lamps. M1, M2 and M3 are grating monochromators; W1, W2, W3 and W4 are compensated circular neutral density wedges. BS: beamsplitter; Mi: mirror; T: fieldstop; Ma: Maxwellian lens; Ch: chopper; Sh: electronic shutter; Ap: artificial pupil. In the right part of the figure the dimensions of the bipartite field of view are shown.

The beams from M2 and M3 were superimposed by means of beamsplitter BS and

the resulting bichromatic beam was projected onto the circular fieldstop T by means of mirror Mi. The monochromatic beam from M1 was also projected onto the fieldstop such that both beams each comprised one half of the field, separated by a thin black line. The Maxwellian lens Ma then focussed the exit slits of the monochromators on the observers pupil, situated behind artificial pupil AP.

Ch is a chopper which was used during flickerphotometric measurements and calibration procedures and Sh is an electronic shutter.

The intensity of each beam could be attenuated by means of compensated circular neutral density wedges (Wratten 98) W1, W2, W3 and W4. Positioning of these filters was achieved through computer controlled stepping motors (Super Electric, M061-FD02) allowing for a transmittance difference of 1% per step and a revolution speed of 4 secs at a total transmittance range of 2 log-units.

The wavelength dials of M1 and M3 were also driven by computer controlled stepping motors (Philips), allowing for a positioning tolerance of ± 0.25 nm.

The field of view was a 1.5 degree circular bipartite field with dark surround; both halves of the field were separated by a vertical black bar, extending 0.4 degrees.

A dental impression byte bar was mounted onto a milling table which could be moved in three dimensions in order to accurately position the observers eye.

Calibration

The wavelength calibrations of the monochromators were performed through

centering various lines of a low-pressure Hg-vapor lamp onto the entrance slits of the monochromators. Half band-widths were 3.4, 3.5 and 3.4 nm for M1, M2 and M3 respectively.

The neutral density wedges were calibrated in situ at 540 nm by means of a photometer (Spectra Pritchard 1980A), focussed at the field stop.

The relative radiance spectra of the sources were also determined in situ by means of the monochromators and the same photometer and using the known (traceable to the U.S. National Bureau of Standards) RSF of the photometer.

For equating radiance units across beams we used the following procedure. Let λ_i denote the wavelength of beam i and λ_j the wavelength of beam j. A photodiode (UDT 450) was placed on the imagepoint of the exit slits of the monochromators. A block-wave (198 Hz) was produced by the chopper Ch in the pathway of the beams and the output signal of the photodiode was analysed by a phase-sensitive detector (Brookdeal Electronics PSD 411) which produced signals r_i and r_j. First, r_i was determined at λ_i and next we determined a position of the density wedge in the other beam j such that $r_i=r_j$ with $\lambda_i=\lambda_j$. Knowing the relative radiance spectra of the sources and the density characteristics of the wedges we were able to compute the radiance ratio of beams i and j at any (λ_i,λ_j)-combination and any position of the wedges.

Procedure

Each function f(s|a,b,q) was determined by having the subject select an intensity t such that, for fixed values of s and q, (2) holds.

The observers were therefore instructed to position wedge W4 such that both halves of the bipartite field appeared equally bright to them.

During all sessions care was taken to select a radiance range of the beam from M1 such that the observer was always able to produce a match without being forced to position the wedge very near the upper or lower limits of the density range of W4. Positioning of W4 was performed through manipulating a joy-stick.

1.0 sec before stimulus onset the subject heard a warning signal during 0.5 sec. At stimulus onset W4 was in a randomly chosen position. The stimulus remained visible during 4 secs.

At stimulus offset the subject decided whether or not the obtained match was satisfactory by pressing either of two knobs. If the observer was not satisfied the same standard was presented again, but now with W4 in the position where it had been set during the previous presentation. If the subject indicated that the obtained match was satisfactory a new level of s (a new position of W1) was chosen. The time interval between two presentations was always 5 secs.

For each $f(s|a,b,q)$ there were six equally spaced log-radiance levels s_i $(i=1,\ldots,6)$ and for each x_i 5 determinations of the corresponding t_{ij} $(j=1,\ldots5)$ were obtained. So, each approximation of $f(s|a,b,q)$ consisted of 30 measurements (s_i,t_{ij}) The order of presentation of different s_i was random for each replication j.

For a we took a fixed wavelength of 550 nm, for b we used 450, 500, 600 and 650 nm and for each wavelength of b we set q, the ratio of radiances of a and b, at 0, 1/3, 1 and 3. So, we approximated 16 different $f(s|a,b,q)$.

Observers

Both authors served as subjects. Both were highly trained and color-normals according to the Munsell-100-hue test.

Results of Experiment I

In Fig. 2[3] the data of subject CW are shown; the figure contains four parts: one for each wavelength of b. The axes represent arbitrary log-radiance units and the means of the $\log(t_{ij})$ are shown along with the 80%-confidence bars. For clarity of presentation the points representing the means have been vertically displaced per value of q, the mixture-ratio.
Under the hypothesis that invariance holds and the assumption that

$$t_{ij} = \nu_{ij} C s_i$$

where C is a constant depending upon a, b and q and ν is a log-normal distributed error-component, we have

$$\log(t_{ij}) = \log(s_i) + \log(C) + \log(\nu_{ij}) \qquad (3)$$

So, in Fig. 2, the means of the $\log(t_{ij})$ should approximately lie on a straight line with a slope of 45 degrees. From the picture it is clear that, given the validity of the statistical assumption, invariance holds for most wavelength-combinations and mixture-ratio's. The validity of the statistical hypothesis is corroborated by the fact that the ratio's of means to variances are approximately constant. Fig. 2 shows the data of subject CW

---

[3]This Figure was also presented in Chapter 2.

only; those of subject CE show the same general pattern since the next analysis leads to similar results for both subjects (see Table 1).

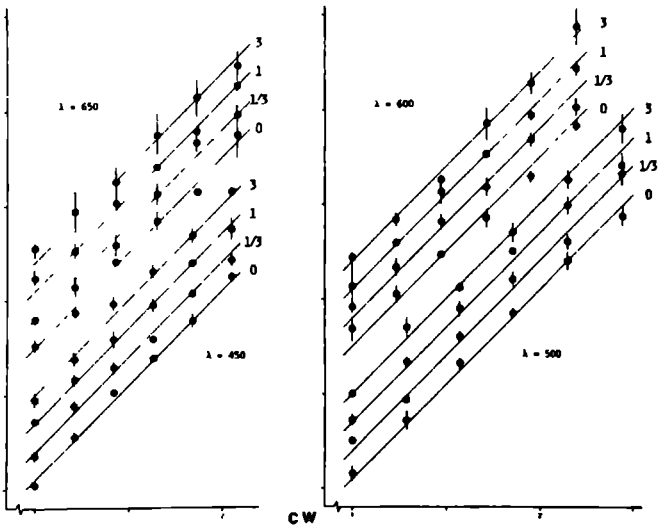


Figure 2. Results of Experiment I for subject CW. The abcissa represents the log-radiance (in arbitrary units) of the reference light of 550 nm; the vertical axis represents log-radiance (in arbitrary units) of the equally bright mixtures of 550 nm and 4 different addend wavelengths. These wavelengths are indicated in the figure. Also indicated in the figure are the four different mixture ratio's $q=R(550)/R(\lambda)$. Points represent the means of 5 replications, the vertical bars denote the 80%-confidence intervals for the means. The drawn lines represent the best-fitting regression lines according to equation (3) (for explanation, see text). The regression lines and the corresponding datapoints have been arbitrarily displaced along the abcissa for clarity of presentation.

A more sophisticated test of invariance was performed using analysis of variance (Winer, 1962, Bonet and Bentler, 1983, see also Note 1 of Elzinga and de Weert, 1984b) per subject per $f(s|a,b,q)$ The results of these analyses are presented in Table 1 The table contains the probability levels p of the F-ratio's from the analyses of variance and the quantities $z=-2\log(p)$ The z may be considered as samples from a χ^2-distribution with df=2 (Johnson and Kotz, 1970) so the sum of these z-values are samples from a χ^2-distribution with df=32 Obviously, for both subjects, these sums are too small to reject invariance of direct brightness matching

Invariance and models

The results of the experiment discussed in the previous paragraphs clearly support the hypothesis that there exists an invariant code for brightness and therefore that the concept of a RSF can be unambiguously attached to direct brightness matching Now the question arises of what restrictions invariance imposes upon models for brightness sensitivity.
Any model for brightness sensitivity that is formulated in terms of action spectra of cone-systems satisfies

$$S_B(\lambda,t) = F\{\rho(\lambda,t)\} \qquad (4)$$

where $S_B(\lambda,t)$ denotes the brightness sensitivity at wavelength λ and radiance t and F an arbitrary function and $\rho(\lambda,t)$ the vector of activities of three cone-systems at (λ,t), i e $\rho(\lambda,t)=\{\rho_1(\lambda,t), \ ,\rho_3(\lambda,t)\}$

Table 1.

Statistics for the results of Experiment I. λ refers to the added wavelengths in the bichromatic mixtures with 550 nm, $p=\text{Prob}(F_{1f} \geq F_{4,24})$ and $z=-2\log(p)$. q denotes the radiance ratio R(550)/R(λ). For explanation see text.

		Subject CW			Subject CE		
λ	q	p	z	p	z	Σz
450	0	0.177	1.504	0.539	0.537	2.077
	1/3	0.905	0.087	0.046	2.674	2.761
	1	0.041	2.774	0.073	2.273	5.047
	3	0.122	1.827	0.381	0.838	2.665
500	0	0.653	0.370	0.010	4.000	4.370
	1/3	0.109	1.925	0.761	0.237	2.162
	1	0.052	2.568	0.678	0.338	2.906
	3	0.709	0.299	0.106	1.949	2.248
600	0	0.672	0.345	0.950	0.045	0.390
	1/3	0.872	0.119	0.547	0.524	0.646
	1	0.821	0.171	0.001	6.000	6.171
	3	0.622	0.412	0.402	0.792	1.104
650	0	0.852	0.139	0.857	0.134	0.273
	1/3	0.559	0.505	0.859	0.096	0.601
	1	0.312	1.012	0.263	1.160	2.172
	3	0.545	0.527	0.347	0.919	1.446
		$\chi^2_{32}=\Sigma z=14.584$		$\chi^2_{32}=\Sigma z=22.574$		37.039

Because of invariance of S_B and the ρ_i we may also write

$$S_B(\lambda,t) = tS_B(\lambda) = tF\{\rho(\lambda)\} \qquad (5)$$

where $S_B(\lambda)=S_B(\lambda,1)$ and $\rho(\lambda)=\rho(\lambda,1)$.

This functional equation was solved by Sirovich (1977) (see also Sirovitch and Abramov, 1977) who assumed that a rational power F^v of F is differentiable in a neighbourhood of zero. Under this hypothesis the solution of the above functional equation is given by

$$F\{\rho(\lambda)\} = \{ \sum_i^3 \alpha_i \rho_i^p(\lambda)\}^{1/p} \qquad (6)$$

i.e. if S_B satisfies invariance and the v-th power of S_B is differentiable in a neighbourhood of zero, brightness sensitivity can be represented with a Minkowski-metric. Obviously, since the basis of the color-matching space is unique up to a linear transformation, Guth's model (e.g. Guth, Donley and Marrocco, 1969) is the special, euclidian case of (6).

However, since S_B exhibits sub-additivity (Ives, 1912c; Kohlrausch, 1935; Dresler, 1953; Tessier and Bloittiau, 1951; Guth, Donley and Marrocco, 1969) and super-additivity (Guth, Donley and Marrocco, 1969; Ingling, Burns and Drum, 1977; Burns, Smith, Pokorny and Elsner, 1982), models of the form of equation (6) should be excluded because Minkowski-metrics either allow for sub- ($p>1$) or super-additivity ($p<1$) but not for both. So the assumption of differentiability of S_B is false as is any assumption leading to (6).

However, without any assumption about F apart from invariance, it is not difficult to derive (Chapter 2, formula 10)

$$S_B(\lambda) = F\{\rho(\lambda)\} = \sum_i^3 \alpha_i \rho_i H\{\rho_2(\lambda)/\rho_1(\lambda), \rho_3(\lambda)/\rho_1(\lambda)\}, \qquad (7)$$

where H is an arbitrary function. Any solution of (5) can be written in the form of equation (7) - although (7) is not the unique solution - and no restriction is imposed with regard to the kind of additivity S_B should satisfy.

The models recently proposed by Breton (1979) and Yaguchi and Ikeda (1983) do not satisfy (7); i.e. do not satisfy invariance, and should therefore be rejected.

Although (7) still leaves us with the problem of specifying H, the expression is not entirely devoid of any interpretation. On the assumption that brightness is not only mediated by an achromatic channel but also by one or more chromatic channels, expression (7) easily lends itself to interpretation: the linear part of (7) could denote the achromatic mechanism and the function H as a combination of chromatic mechanisms.

This interpretation is entirely within the mainstream of modern concepts of the brightness mechanism and, by its formulation, closely resembles a widely neglected proposal of Wasserman (1969; Wasserman and Gillman, 1970).

Wasserman proposed

$$S_B(\lambda,t) = k_1 A(\lambda,t) + k_2\{C(\lambda,t)\}^{k_3} \qquad (8)$$

where $A(\lambda,t)$ denotes the (nonlinear) achromatic channel and $C(\lambda,t)$ the (linear) difference between two (nonlinear) chromatic channels. The k_i denote constants of which $k_3>1$ (Wasserman chooses $k_3=1.5$). Wasserman showed that a model like (8) is capable of producing both super- and sub-additivity and the superadditive effect of chromatic backgrounds on S_B (Boynton and Das, 1966; Boynton, Das and Gardiner, 1966).

However, just like the models of Breton and Yaguchi and Ikeda, the model

does not satisfy invariance.

A new model for brightness sensitivity

As a specification of (7) that both satisfies invariance and appropriately behaves super- and sub-additively we propose

$$S_B(\lambda) = \sum_i^3 \alpha_i \rho_i(\lambda)\{\beta_1 \sigma(\lambda) + \beta_2 \mu(\lambda)\}^{\gamma}, \qquad (9)$$

$$\sigma(\lambda) = \rho_2(\lambda)/\rho_1(\lambda),$$

$$\mu(\lambda) = \rho_3(\lambda)/\rho_2(\lambda),$$

where the α_i, β_i and γ are constants, all positive except for γ which we assume to lie in the interval (0,-1). In the above equation the ρ_i (i=1,2,3) denote action spectra of cone systems with peak sensitivities in the long, middle and short wave regions.

Note that σ and μ do satisfy invariance but are nonadditive since

$$\rho_i(a \odot b)/\rho_j(a \odot b) < \rho_i(a)/\rho_j(a) + \rho_i(b)/\rho_j(b)$$

A model like (9) should not only generate more or less realistic super- and sub-additivities for bichromatic mixtures but it should also be capable of generating appropriate direct brightness matching sensitivity functions and it should describe brightness sensitivity in the case of monochromatic mixtures with a white light as measured by Burns et al. (1982).

To show the potential power of the model we will therefore estimate its parameters from empirical sensitivity functions and simulate the results of some relevant experiments with the estimated parameters. Experiment II

provides us with the necessary spectral sensitivity data

Experiment II

The experiment to be presented below is by no means original it merely consists of determining sensitivity functions for direct brightness matching for the subjects that also served in Experiment I Formally, the experiment amounts to determining t in aE(t*b) where the a are monochromatic lights from an equal-energy spectrum and b a monochromatic reference light of 550 nm

The purpose of the experiment is just to provide us with data that are sufficient to estimate (ratio's of) the parameters of our model and determining the fit of the model to several kinds of relevant data

Apparatus and Calibration

The apparatus used was the same as in Experiment I, except for the fact that density wedge W4 was placed in the beam of monochromator M1 that provided the monochromatic reference light of 550 nm in the right half of the bipartite field The beam from M3 provided the monochromatic equal-energy stimuli. The field of view was the same as in Experiment I as were the calibration procedures.

Procedure

The stimuli were monochromatic lights with wavelengths ranging from 420 up

to 680 nm in steps of 10 nm, presented in random order. The stimulus presentation procedure and the subjects task were the same as in Experiment I. The subject adjusted the intensity of the beam from M1 by adjusting the position of W4. At each wavelength from M3 5 measurements of t in (10) were made.

Results of Experiment II.

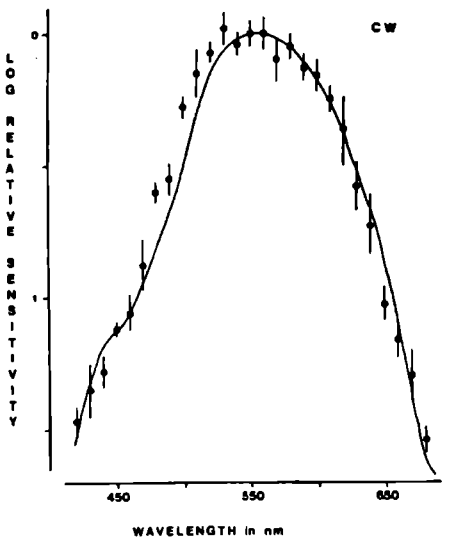


Figure 3. Results of Experiment II for subject CW. Log relative sensitivity for direct brightness matching, normalized to $\log(S_B(550))=0$. The points represent the means of 5 replications; the vertical bars represent 80%-confidence intervals. The solid line represents log relative sensitivity according to equation (9) (for explanation, see text) and the corresponding parameter estimates for subject CW.

The means of the values of $\log(t_{ij})$ are shown in Fig 3 for subject CW together with the 80%-confidence bars at each wavelength presented (normalized at 550 nm) Similar results were obtained for subject CE The results are very much like the results published by Wilson (1964) and Comerford and Kaiser (1975) Note that t was also determined for a=b in (9) and that the variance of these metameric matches is not the smallest

Parameter estimation

In terms of our model equation (9) reads

$$\sum_{i}^{3}\alpha_i\rho_i(550)\{\beta_1\sigma(550)+\beta_2\mu(550)\}^{\gamma} - t\sum_{i}^{3}\alpha_i\rho_i(\lambda)\{\beta_1\sigma(\lambda)+\beta_2\mu(\lambda)\}^{\gamma} = 0 \qquad (10)$$

Once a choice has been made regarding the exact shapes of the ρ_i and the parameter values are given, t can be solved for

For the ρ_i we choose, quite arbitrary, the R, G and B cone sensitivities as proposed by Vos (1978a,b) The parameters were then estimated by minimizing the function

$$\zeta = \sum_{i}(t_i-t_i^*)^2/t_i$$

where the t_i are the means of the settings obtained in Experiment II and the t_i^* denote the solutions of equation (10) for given parameter estimates Actual minimization of ζ was performed through the pattern-search algorithm of Hooke and Jeeves (1961)

However, not all the parameters appearing in equation (10) are identifiable on the basis of our data from Experiment II since one of the α_i and one of the β_i can be divided out from equation (10) Arbitrarily, we divided

equation (10) by $\alpha_2\beta_2^{\gamma}$ so the only parameter that can be directly estimated from our data is γ, the other estimates are ratio's of parameters. We will, however, not carry this through in our subsequent notation but merely adopt the convention $\alpha_2=\beta_2=1$.

Note that the constants in equation (10) are justified by the experimental procedure but for the estimation procedure it is completely immaterial whether 550 nm or any other wavelength is chosen as the reference wavelength.

The results of the estimation procedure are tabulated in Table 2 for both subjects.

Table 2.

Estimates of the parameters of equation (10) per subject

	CW	CE
ζ	0.215	0.325
α_1	0.211	0.047
α_3	5.019	4.165
β_1	0.2363	0.0041
γ	-0.396	-0.819

The drawn line in Fig. 3 represents the RSF generated with the model and the parameter estimates for subject CW. Obviously, the fit of the model does not seem to be too bad at all.

Next, we consider the capability of the model to generate realistic super- and sub-additivity of brightness sensitivity. To this end we first define

the model-independent additivity-index $A_B(a,b|q)$ as

$$A_B(a,b|q) = S_B\{q*a \oplus (1-q)t*b\}/S_B(a)$$

$$= S_B\{q*a \oplus (1-q)t*b\}/S_B(t*b)$$

with t such that $S_B(t*b)=S_B(a)$

Note that if S_B is super-additive, $A_B>1$ and if $A_B<1$ S_B is sub-additive, in the case additivity holds, $A_B=1$.

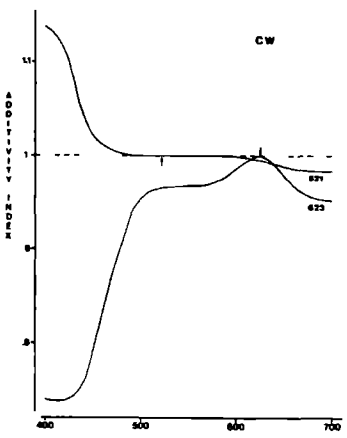


Figure 4 Simulated additivity indices with the parameter estimates of subject CW and equation (9) for bichromatic mixtures of the wavelengths shown on the abcissa and 521 and 623 nm (arrows)

This additivity-index is much more informative about brightness matching than indices that completely or partly take flickerphotometric luminances

into account, such as the B/F-ratio as used by Burns et al. (1982) or the index used by Ikeda (1983) and Yaguchi and Ikeda (1983).

Given the estimates of the parameters of our model we can generate A_B-values. In Fig. 4 the values of model-generated $A_B(521,\lambda|0.5)$ and $A_B(623,\lambda|0.5)$ have been plotted for λ varying between 400 and 700 nm. The A_B were computed using the parameter estimates obtained for subject CW. We chosed these particular A_B because 521 and 623 nm are the fixed wavelengths

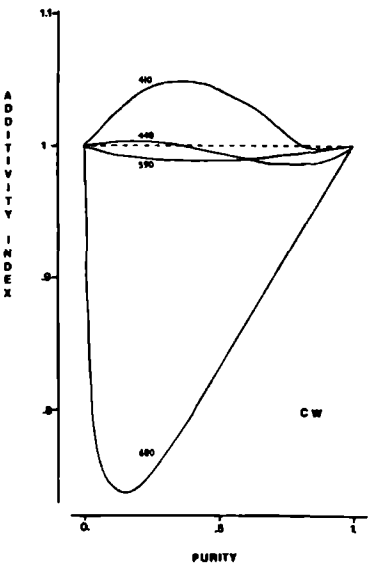


Figure 5. Simulated additivity indices for subject CW for mixtures of varying purity of white light (x=0.362, y=0.394) with the wavelengths indicated in the figure.

at which Guth, Donley and Marrocco (1969, experiment 5) determined supra-

threshold brightness matches for bichromatic mixtures. The pattern of the model-generated super- and sub-additivities is qualitatively equal to that found by Guth et al. (1969). For subject CE very similar results were obtained. Burns, Smith, Pokorny and Elsner (1982) found brightness super-additivity for short wave mixtures of white (x=0.362, y=0.394) of varying purity. We simulated this experiment also: in Fig. 5 A_B has been plotted for various wavelengths and purities on the basis of the parameter estimates of subject CW. Again the simulation results are consistent with the data of Burns et al. (1982) and similar results were obtained with the parameter estimates for subject CE.

Note that for short wavelength addends the kind of additivity depends on purity; such transients were also observed by Burns et al.

Sensitivity and magnitude estimation

In the introduction we already mentioned a number of authors whose findings indicate that brightness magnitude estimates behave as a power function of intensity:

$$\psi_B(\lambda,t) = a(\lambda)t^b$$

$a(\lambda)$ is a wavelength dependent constant and b a constant that is independent of wavelength, approximately equal to 0.33.

Assuming that

$$\psi_B(\lambda,t) = H\{tS_B(\lambda)\} \qquad (11)$$

with $H(x)=kx^{0.33}$, leads to

$$a(\lambda) \approx k S_B(\lambda)^{0.33}$$

Ekman, Eisler and Künnapas (1960) estimated $a(\lambda)$ for 8 different wavelengths in the range from 469 up to 672 nm. In Fig. 6 we plotted $S_B(\lambda)^{0.33}$ against wavelength on the basis of subject CW's parameter estimates. Also shown in Fig 6 are the estimates of $a(\lambda)$ as provided by Ekman et al. (1960) and the spectral sensitivity curves obtained by Cavonius and Hiltz (1973) through brightness magnitude estimation.

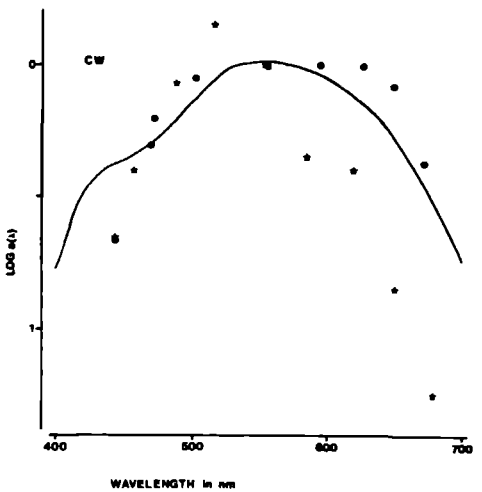


Figure 6. The drawn line represents the logarithm of $a(\lambda)$ appearing in $\psi_B(\lambda,t)=a(\lambda)t^{1/3}$ simulated with (9) and the parameter estimates obtained for subject CW The ★ represent log{$a(\lambda)$} as determined by Cavonius and Hiltz (1973), the ● represent log{$a(\lambda)$} as determined by Ekman et al.(1960)

Although the paradigm used by Ekman et al is identical to the paradigm used by Cavonius and Hiltz, their results are quite different in a strict, quantitative sense. Qualitatively, the results agree: sensitivity, measured through magnitude estimation, peaks at middle-wavelengths and is low at the extremes of the spectrum. As shown in Fig. 6, this is also predicted by equation (11). Considering the substantial quantitative differences between the results of these two identical experiments, one should not expect any model to do a much better job in predicting these results than our model does.

A second indication that (11) might be valid is provided by the results of Elzinga and de Weert (1978). They obtained brightness estimates for bichromatic mixtures of yellow and green. It appeared that, for intermediate levels of admixed green, the brightness magnitude estimates show a tendency to decrease with increasing intensity of the yellowish wavelength down to some minimum and than to increase with further increasing intensity. This is exactly what is to be expected given (11) and the sub-additivity of yellow-green mixtures.

Discussion

We feel that the most salient aspect of the material presented in this paper is the demonstrated fact that the results of a Class B experiment (Brindley, 1960b) provides us with a sensitivity function that is uniquely determined up to multiplication by a positive constant.

The second important point we have tried to make is that invariance, although not sufficient to derive metric representability of brightness

sensitivity, leads to a restriction on possible models for S_B that has heuristic potentials. This is shown by the results of the simulation of various kinds of experimental data. We do, however, not pretend that the model we proposed is correct in a very strict, quantitative sense; we think that, at the present time, it is not very useful to try to improve upon the quantitative fit of models beyond the point we reached since that would lead to mere curve fitting because theories on the structure of the mechanism that generates the sensation of brightness are too limited and the available data base is too small.

Often, but not always, the empirical brightness sensitivity curve shows a local minimum beyond 550 nm. Our model cannot generate such a local minimum. This is a qualitative flaw. However, the model satifies invariance, does provide a reasonable description of the sensitivity function and simulates super- and sub-additivity, so we are are not willing to reject it for this reason only.

The last point to be understood clearly is that invariance of direct brightness matching implies that direct brightness matching data are not affected by the intensity level at which they are obtained. The fact that detectability-data indicate that super-additivity does not occur at threshold level is thus strong evidence against the conjecture of Guth et al. (1969) that a detection experiment is the limiting case of supra-threshold direct brightness matching. Thus, models for threshold data cannot be valid for supra-threshold experiments on brightness perception. Again, this was recognized earlier by Wasserman (Wasserman and Gillman, 1970) who had to make an extra assumption concerning the behavior of chromatic channels at threshold-intensities in order to make their model generate

acceptable sub-additivity at threshold.

chapter 4:

NONLINEAR CODES FOR THE YELLOW/BLUE MECHANISM[1]

Abstract

The opponent yellow/blue mechanism is studied by means of an iso-cancellation technique. A qualitative and quantitative analysis of different photopigment models for the yellow/blue code is presented. On the basis of this analysis we provide a new model which ascribes the nonlinear character of the code to a power transformation of the short wave cone activity. This new model leads to predictions concerning the wavelength shift of the short wave component of a unique red. Tests of these predictions support the new model.

Introduction

Opponent colors theory as first formulated by Hering (1878) assumes the existence of two separate opponent chromatic channels.

Jameson and Hurvich (1955) first tried to measure the sensitivity functions of these mechanisms and called them cancellation functions because of the nature of the procedure through which these measurements were obtained.

To measure the sensitivity of, for example, the yellow/blue channel for equally luminant spectral lights, Jameson and Hurvich measured the radiance of an admixed yellow that is required to cancel the sensation of blueness. For yellowish spectral lights the radiance of an admixed blue required to cancel the perceived yellowness was measured. By transforming their measurements to an equal energy spectrum, they implicitly assumed that the opponent chromatic channels could be adequately described by a linear

---

[1]C.H. Elzinga & Ch.M.M. de Weert (1984) Nonlinear codes for the yellow blue mechanism. *Vision Research, 24, 9,* 911-922.

combination of receptor absorption spectra and thus that their so obtained sensitivity functions were invariant over different energy levels.

Krantz (1975a,b) formulated a measurement theory that explicitly stated the conditions under which the transformation from an equal energy to an equal luminance spectrum is justified. These conditions amount to the requirement that the set of lights that produce a state of equilibrium in one or both of the opponent chromatic mechanisms is closed under additive and multiplicative color mixture operations.

Several investigators (Larimer, Krantz and Cicerone, 1974, 1975; Cicerone, Krantz and Larimer, 1975; Raaijmakers & De Weert, 1975; Moeller, 1976; Nagy, 1979; Ikeda and Ayama, 1980) have tested these conditions. On the basis of these investigations it is now generally accepted that closure holds indeed for the red/green mechanism but not so for the yellow/blue system (but see Savoie, 1973; Cornsweet, 1978).

Since the closure properties hold for the red/green mechanism the sensitivity function of that mechanism is a linear function of colorimetric primaries.

Formally:

$$\phi_1(\lambda,t) = x_1 t\alpha(\lambda) + x_2 t\beta(\lambda) + x_3 t\gamma(\lambda), \qquad (1)$$

where $\phi_1(\lambda,t)$ denotes the sensitivity of the opponent red/green mechanism at wavelength λ and radiance t; α, β and γ denote the spectral sensitivities of three independent cone systems with peak sensitivities in the short, middle and long wave region respectively. Finally the x_1 represent constants, not all having the same sign. If we denote the sensitivity of the opponent yellow/blue system as $\phi_2(\lambda,t)$ we could also write:

$$\phi_2(\lambda,t) = F\{t\alpha(\lambda),t\beta(\lambda),t\gamma(\lambda)\} \qquad (2)$$

Since the closure properties are violated for the yellow/blue system the function F in (2) cannot be linear (but see Guth, Massof and Benzschawel, 1980).

Several investigators (Larimer, Krantz and Cicerone, 1975; Werner & Wooten, 1979a,b) have proposed different nonlinear forms of F in (2). Below we will discuss these models in detail and we will show that none of these models is able to explain all existing data on closure.

We felt, however, that the available data that indicate nonlinearity of F do not reveal the precise nature of this nonlinearity. Since the nonlinearity of F depends upon the way closure is violated for the yellow/blue mechanism we tested multiplicative closure in some detail.

Formally, multiplicative closure amounts to the assumption

$$a \epsilon C_2 \quad \text{iff} \quad t*a \epsilon C_2 \text{ for all } t>0. \qquad (3)$$

Here C_2 denotes the set of all lights a that produce a state of equilibrium in the yellow/blue channel i.e. these lights look neither yellowish nor bluish, and the operation * denotes a multiplication of the radiance by some positive number t.

It is well known that, except for wavelengths around 510 nm (a so-called equilibrium green), the hues of all monochromatic yellowish-green or bluish-green lights varies with intensity (the Bezold-Brücke hue shift).

On the other hand, for all mixtures of a wavelength shorter than η and a wavelength longer than η there exists a radiance ratio such that the mixture belongs to C_2: for all triples λ,s,μ with $\lambda>\eta$ and $\mu<\eta$ there exists a $t>0$ such that the bichromatic mixture $(\lambda,s;\mu,t) \epsilon C_2$. The first experiment to be

discussed here amounts to the measurement of t for different triples λ,s,μ For fixed wavelength-pairs λ,μ this results in the determination of points on the graph of the function $f_{\lambda,\mu}(s) = t$. Such a function is called an iso-cancellation function.

If multiplicative closure (3) were to hold for C_2, then the graphs of all iso-cancellation functions would be straight lines with the intercept at the origin (it is understood that s and t have the same dimension).

Any systematic violation of multiplicative closure would be revealed by a systematic deviation of those graphs from a straight line A sufficiently fine grid of points on such a graph should lead to a close approximation of its shape. Knowledge of the properties of iso-cancellation functions could lead to understanding the nonlinear character of the yellow-blue mechanism and the specification of a nonlinear model in terms of action-spectra of cone systems.

Methods

a) Apparatus

The stimuli consisted of a mixture of two monochromatic beams, optically superimposed by a beamsplitter. (The apparatus used is part of a more complicated 3-channel Maxwellian view system. We will only describe the components used in the experiments reported in this paper.) The foveally presented field of view consisted of a circle subtending 1.4°; except for this circle the field of view was dark.

The monochromatic beams are produced by two Oriel grating monochromators. Wavelength calibrations were performed by centering various lines of a low-pressure Hg vapor lamp on the entrance slits of the monochromators M1 and

M2. Halfband widths were 3.4 nm for M1 and 3.5 nm for M2. The wavelength dial of M1 is driven by a computer-controlled steppingmotor (Philips), allowing for a positioning tolerance of 2.5 nm

The light sources S1, S2 and S4 were 220V, 150W halogen tungsten ribbon filament lamps (underrun at 0.6A, AC-stabilized voltage). Light from M1 and M2 was mixed by beamsplitter BS1. Light source S4 was used for flickerphotometric measurements in combination with chopper Ch rotating behind field-stop T. S3 is a red LED and produced an optical warning signal 0.5 sec. before each stimulus presentation. The image of S3 was projected 4° temporal from the stimulus by means of beamsplitter BS2.

The intensity of the beams from M1 and M2 was controlled by two circular compensated neutral density wedges (Wratten 98) W1 and W2. Positioning of W1 and W2 was performed by means of computer controlled stepping motors (Super Electric, M061-FD02), allowing for a transmittance difference of about 1% per step and a revolution speed of 4 secs at a total transmittance range of 2 log-units.

The beams from M1 and M2 are superimposed by means of a beamsplitter BS1 and the resulting beam is projected onto the field stop T. The Maxwellian lens M1 focused the exit slits of M1 and M2 and the LED on the observers pupil, situated behind an artificial pupil AP.

A dental impression byte bar was mounted to a milling table which could be moved in three dimensions in order to accurately position the observers eye.

b) Calibration

The neutral density wedges were calibrated in situ at 540 nm by means of a photometer (Spectra Prittchard, 1980A) focused on the field stop.

The relative radiance spectrum of the beam from M2 was determined by means of the same photometer focussed on the field stop and using the known relative sensitivity of the photometer

For equating radiance units across beams from M1 and M2 we used the following procedure Let λ denote the wavelength of the beam from M1 and μ the wavelength of the beam from M2 A photodiode (UDT 450) was placed on the imagepoint of the exit slits of M1 and M2 A block-wave (198 Hz) was produced by chopper Ch in the pathway of the beams and the output signal of the photodiode was analysed by means of a phase-sensitive detector (Brookdeal Electronics PSD 411), which produced signals r1 and r2 for beams from M1 and M2 respectively

First r1 was determined at λ nm Next, we determined a position of W2 such that r2=r1 with the beam from M2 at λ nm also Having determined the relative radiance spectrum of the beam from M2 and the density characteristics of W1 and W2 we were able to compute the radiance ratio of the two beams at any (λ,μ)-combination and any position of the wedges W1 and W2.

Experiment I Multiplicative closure of C_2

Procedure

Each isocancellation function $f_{\lambda,\mu}(s)$ was determined by having the subject select an intensity t such that the foveally presented pair $(\lambda,s,\mu,t)\epsilon C_2$ The observers were instructed to turn a neutral density wedge in the path of the μ-beam, increasing or decreasing the radiance until the resulting stimulus appeared neither yellowish nor bluish

Each observer had one practice session for each λ,μ-pair prior to data

collection.

During all the sessions we selected a radiance range of the λ (yellowish)-beam such that the observer was able to position the wedge in the path of the μ (bluish)-beam such that the resulting stimulus could be made to appear both too yellowish and too bluish. 1.0 sec before stimulus onset the observer saw a LED-produced warning signal during 0.5 sec. At stimulus onset the stimulus was composed of a fixed radiance s_i at λ nm and a randomly chosen radiance t_0 at μ nm. The stimulus remained visible during 4.0 secs and during that time the subject had to position the μ-wedge by means of a joy-stick. He was instructed to produce a position of the wedge such that the resulting stimulus appeared neither yellowish nor bluish. At stimulus offset the subject decided whether the obtained wedge-position was satisfactory or not and communicated his decision by pressing one of two knobs. If the observer was not satisfied, the same mixture was presented again, but now with the μ-wedge in the position where it was set during the previous presentation. If the subject indicated having met the criterion a new mixture was presented. The time between two presentations was always 6.0 secs.

For each $f_{\lambda,\mu}(s)$ there were 10 equally spaced radiance levels s_i (i=1,..,10) and for each radiance level s_i 10 measurements of the corresponding t_{ij} (j=1,..,10) were obtained. So the approximations of the $f_{\lambda,\mu}(s)$ consisted of 100 points (s_i,t_{ij}). The order in which the s_i were presented was random for each replication j.

Observers

Three male observers took part in the experiment. All three were color

normals One of the subjects, HE, had no prior experience in color vision experiments, the other two observers were highly experienced.

Results of Experiment I

Since each observer produced definite outliers in most combinations (s_i, t_{ij}) we decided to remove the two most extreme values of t_{ij} for each value of s_i for all (λ,μ)-combinations All further analyses reported in this paper are based on the so trimmed datasets, thus consisting of 8 replications per s_i-value per subject.

The means of these replications are depicted graphically in Fig. 1 for each observer separately.

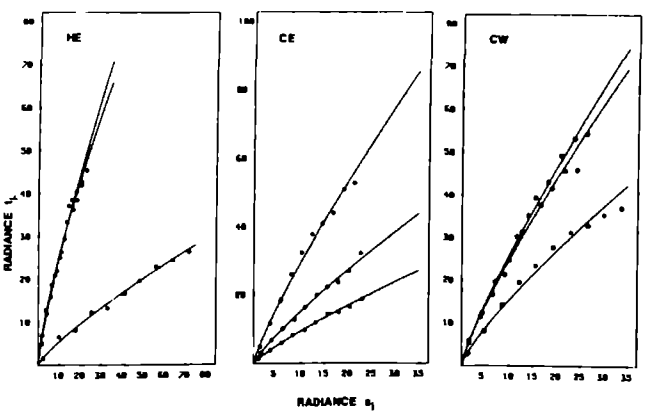


Figure 1. Results of Experiment I. Isocancellation curves from experiment I per subject per wavelength combination. The points locate the means of 8 measurements, the lines represent the curves according to $t=\nu Bs^C$ (for explanation see text). The radiance units of s and t are arbitrary but equal

In order to test for multiplicative closure of C_2 we estimated the constant A in the straight line, zero-intercept equation

$$t_{ij} = As_i + \varepsilon_{ij}$$

separately for each observer and each λ,μ-combination with the appropriate least squares estimator. Analysis of variance (Winer, 1962; see also Note 1) indicates that this regression model is probably incorrect for all observers and all λ,μ-combinations. We therefore conclude that systematic violations of multiplicative closure occur for all observers and all three wavelength-pairs.

Table I.

Summary of the results of fitting $t_{ij}=As_i+\varepsilon_{ij}$ per wavelength combination per subject. $a=\Sigma s_i^2/\Sigma t_{ij}s_i$ which is the least squares estimate of A. F_{lf} is the F-ratio associated with lack of fit of the model (see Note 1), η_m^2 and η_{lf}^2 denote the correlation ratio's associated with the model and lack of fit respectively. r^2 denotes the squared linear correlation between data and the predictions generated from the model.

	(585,495)			(570,480)			(555,465)		
	HE	CE	CW	HE	CE	CW	HE	CE	CW
a	2.31	2.75	2.32	2.26	1.42	2.21	0.40	0.86	1.28
F_{lf}	1.73	4.85	6.00	6.58	1.20	6.84	0.66	2.73	22.23
η_m^2	0.862	0.946	0.962	0.942	0.947	0.935	0.854	0.960	0.947
η_{lf}^2	0.023	0.019	0.015	0.025	0.006	0.029	0.010	0.009	0.038
r^2	0.974	0.980	0.984	0.974	0.993	0.970	0.988	0.990	0.962

From Fig. 1 it appears that the iso-cancellation functions are concave. It

appeared that the ratio of the means of the t_{ij} to the variance around those means is approximately constant.

We therefore estimated the constants B and C in the model

$$t_{ij} = \nu_{ij} B s_i^C$$

where ν denotes a lognormal distributed error component (Heien, 1968). Analysis of variance indicates that this nonlinear model performs better than the linear zero-intercept model does, both in terms of the obtained F-ratio's and in terms of the squared correlations r^2. The main results of this analysis are reported in Table II, together with the estimates of B and C per subject per wavelength pair.

Table II.

Summary of the results of fitting $t_{ij}=\nu_{ij} B s_i^C$ per wavelength combination per subject. b and c denote the unbiased least squares estimators of B and C respectively.

	(585,495)			(570,480)			(555,465)		
	HE	CE	CW	HE	CE	CW	HE	CE	CW
b	4.60	4.18	3.77	3.85	2.15	3.79	0.85	1.33	2.40
c	0.75	0.85	0.84	0.82	0.85	0.82	0.81	0.85	0.81
F_{lf}	1.25	2.98	3.81	4.03	1.37	4.62	0.47	1.54	12.37
η_m^2	0.871	0.955	0.969	0.953	0.947	0.946	0.858	0.956	0.967
η_{lf}^2	0.014	0.010	0.009	0.013	0.006	0.017	0.006	0.005	0.018
r^2	0.984	0.989	0.991	0.986	0.993	0.982	0.993	0.995	0.981

Because of these results we conclude that multiplicative closure is violated for C_2-lights and thus that no adequate linear model for ϕ_2 exists[2]. Note that, within observers, the estimate of C is approximately constant over wavelength pairs, that interobserver variability is small and that, because of the calibration procedure, the estimate of C is dimensionless.

Another salient feature of the analysis is that it seems that the way the estimates of B depend upon the particular wavelength combination is not much different for these observers.

Data and nonlinear models

Below we shall discuss a number of specifications of F in (2), i.e. a number of nonlinear models for ϕ_2 in terms of action spectra of cone systems, together with their database.

Larimer, Krantz and Cicerone (1975) tested multiplicative closure of C_2 by testing wavelength invariance of C_2-lights, i.e. wavelength invariance of unique green (around 510 nm) and unique red (a mixture of 650 nm and a short wave component).

All four observers in the Larimer et al. study produced a significant wavelength shift of the short wave component of a unique red when the total radiance of a mixture of 650 nm and a variable short wave component with a constant radiance ratio ($\approx$ 3.4) was increased. This result (a similar result was obtained by Nagy, 1979) was the main reason for their specification of F in (2):

---

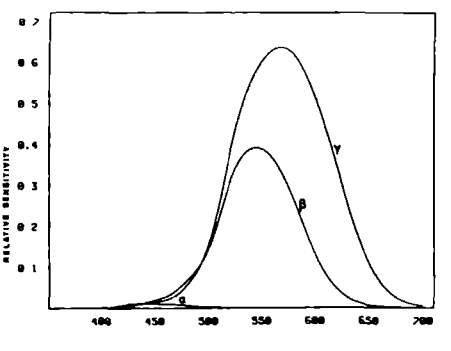
[2]The upper bounds of the 80-% confidence intervals for C are smaller than 1.0 for each subject and each wavelength-combination.

$$\phi_2(\lambda,t) = x_1 t\alpha(\lambda) + x_2 t\beta(\lambda) + x_3\{t\gamma(\lambda)\}^{x_4} \qquad (4)$$

with the parameter $x_4<1$. Such a model indeed predicts the observed wavelength shift of the short wave component in a unique red because the negative acceleration of the activity of the long-wave sensitive part can be compensated for by decreasing the quantum catch of the α-system through increasing the wavelength of the short-wave component.

Larimer et al. estimated the parameter x_4 and the ratio's x_2/x_1 and x_3/x_1 by taking for α, β and γ the cone system action spectra as proposed by Vos and Walraven (1971).

For all observers it turned out that indeed $x_4<1$ and that the fit of the model was reasonably good. Below we plot[3] the Vos-Walraven primaries, as modified by Vos (1978a), for reference.



Relative spectral sensitivity functions of the cone systems as proposed by Vos (1978a). The curves have been normalized such that $\alpha(\lambda)+\beta(\lambda)+\gamma(\lambda)=y(\lambda)$ where $y(\lambda)$ represents Judd's (1951) luminousity function as modified by Stiles (1955).

---

[3]These plots do not occur in Elzinga and de Weert (1984b) but were added for the clarity of the argument.

From the above figure and the model (4) it is evident that this model will also predict a wavelength shift of unique green in the direction of longer wavelengths.

For two observers Larimer et al. found a significant wavelength shift for unique green in the direction of longer wavelengths when increasing radiance (we will treat this phenomenon in the last paragraph of this paper); the other two of their observers did not produce such a shift. Larimer et al. tried to predict their observations on unique green by means of model (4) and the best fitting set of parameters. They found that the model predicted a non-observed wavelength shift of unique green for two of the four observers. They therefore concluded that the model must be qualitatively wrong and instead proposed

$$\phi_2(\lambda,t) = x_1 t\alpha(\lambda) + x_2 t\beta(\lambda) + \delta x_3 |t\gamma(\lambda) - t\beta(\lambda)|^{x_4} \qquad (5)$$

with $x_4<1$ and $\delta=\text{sign}\{t\gamma(\lambda)-t\beta(\lambda)\}$.

Larimer et al. reasoned that this model could both incorporate the wavelength shift of the short wave component in a unique red mixture because $\gamma(\lambda)-\beta(\lambda)$ approaches $\gamma(\lambda)$ when $\lambda>640$ and also, depending on the exact parameter values, either no or just a slight wavelength shift of unique green since, around 510 nm, $\gamma(\lambda)-\beta(\lambda)$ is very small. Again they estimated parameters and predicted unique green loci for different radiance levels. Quantitatively, the fit of model (5) was comparable to that of model (4). In a qualitative sense it was, however, not better than model (4) because it failed to predict the observed wavelength shift of unique green for two observers.

Werner and Wooten (1979a,b) determined a yellow/blue cancellation function for the entire spectrum between 400 and 700 nm at just one level of constant luminance (!) for three observers They compared the fit of a linear model with the fit of model (5) and found that (5) did a considerably better job than the linear model

However, model (5) systematically underestimated the empirical cancellation functions around 580 nm Therefore they proposed the model (see also Note 2)

$$\phi_2(\lambda,t) = x_1 t\alpha(\lambda) + \delta|x_2 t\beta(\lambda) + x_3 t\gamma(\lambda)|^{x_4} \qquad (6)$$

where δ=sign$\{x_2 t\beta(\lambda) + x_3 t\gamma(\lambda)\}$

Werner et al claim that this model produces the best fit to their data and to data from Jameson and Hurvich (1955) and Romeskie (1978) For all subjects it appeared that $x_4 > 1\ 0$ ($1\ 06 < x_4 < 4\ 39$) However, the plots of their data together with the model generated cancellation functions show a systematic displacement of the predicted curve in the direction of shorter wavelengths

Werner and Wooten estimated model-parameters by taking iodopsin-nomograms for α, β and γ They also report (Werner and wooten, 1979a) having used a modification of the Vos-Walraven primaries (Vos, 1978a) which gave essentially the same results We therefore feel safe to compare all the models discussed in the present paper on the basis of this modified version of the Vos-Walraven primaries

The database has now been enlarged by the concave iso-cancellation functions discussed in this paper Note that this concavity is in qualitative agreement with the concavity of the results of increment threshold cancellation experiments reported by Pugh and Larimer (1980)

This extended database may now be used to qualitatively evaluate the nonlinear models discussed so far

Qualitative analysis

In this paragraph we will evaluate the models discussed with respect to their ability to account for the relevant empirical phenomena The models should account for

i) wavelength invariance of unique green
ii) non-existence of a spectral unique red
iii) concave iso-cancellation functions

The first model proposed by Larimer et al (equation 4) should be rejected because it cannot meet the first requirement, i e it cannot generate wavelength invariance of unique green

The qualitative analysis of the remaining two models is more involved The general procedure employed is that we will derive certain restrictions on the parameters from the first two requirements and then confront the models with the third requirement

If, finally, the third requirement cannot be met, we conclude that there is at least one kind of relevant data that the model does not account for, so that model should be rejected If, on the other hand, the model does meet the third requirement given the parameter restrictions, a quantitative analysis of the model is indicated

Furthermore we will adopt the (arbitrary) scaling convention that in each model we take the multiplier attached to the α-system equal to -1, i e x_1=-1 Any other convention would not affect the conclusions of the

analysis. We wil start with the model of Werner and Wooten (equation 6).

First suppose the model generates a wavelength invariant unique green in the region between 500 and 520 nm. This could only be accomplished if the contribution of the α-system is negligeable in this region, the parameters x_2 and x_3 having opposite signs and their absolute values not too much different, i.e. $|x_2| \approx |x_3|$. Because of the non-existence of a spectral unique red and our scaling convention ϕ_2 is positive beyond 520 nm so we must require $x_3 > 0$.

With these restrictions on the parameters in mind we confront equation 6 with the third requirement, the concavity of iso-cancellation functions. Formally, an iso-cancellation setting in terms of equation 6 corresponds to

$$-\{t\alpha(\mu)+s\alpha(\lambda)\}+\delta|x_2\{t\beta(\mu)+s\beta(\lambda)\}+x_3\{t\gamma(\mu)+s\gamma(\lambda)\}|^{x_4} = 0$$

Since in our experiments $\lambda > 550$ nm we shall take $\alpha(\lambda)=0$. Obviously $\delta=1$. The above equation is an implicit form of the iso-cancellation function $f(s)=t$:

$$g(s,t) = 0 = -ut + (pt + qs)^{x_4}$$

$$u = \alpha(\mu)$$

$$p = x_2\beta(\mu) + x_3\gamma(\mu)$$

$$q = x_2\beta(\lambda) + x_3\gamma(\lambda)$$

The requirement that f(s) is monotone and concave formally corresponds to

$$f'(s) = df/ds > 0$$

$$f''(s) = df'/ds < 0$$

These derivatives can be obtained by implicit differentiation with respect to the parameters of g:

$$f'(s) = -g_s/g_t$$

$$f''(s) = -(g_{ss}/g_t) + (2g_s g_{st}/g_t^2) - (g_s^2 g_{tt}/g_t^3)$$

and

$$g_s = x_4(pt+qs)^{x_4-1} q$$

$$g_t = -u + x_4(pt+qs)^{x_4-1} p$$

$$g_{ss} = x_4(x_4-1)(pt+qs)^{x_4-2} q^2$$

$$g_{st} = x_4(x_4-1)(pt+qs)^{x_4-2} pq = g_{ts}$$

$$g_{tt} = x_4(x_4-1)(pt+qs)^{x_4-2} p^2$$

We have to consider the sign of f''(s) so we have to determine the sign of the above partial derivatives.

First, because $\lambda>550$ nm, $x_3>0$ and $|x_2|\approx x_3$ we must have $q>0$ and hence $g_s>0$. We take $f'(s)>0$ and now, because $g_s>0$, we derive $g_t<0$. Following Werner and Wooten we take $x_4-1>0$ so $g_{ss}>0$ and $g_{tt}>0$.

Now the sign of f''(s) critically depends upon the sign of g_{st} which in turn depends upon the sign of p. The sign of p will be positive in the region beyond 490 nm since beyond that point $\gamma(\mu)>\beta(\mu)$.

It is not difficult to see that whenever $p>0$ the iso-cancellation function is definitely convex.

Note that in our experiments one of the iso-cancellation functions was determined at μ=495 nm and that that curve was definitely concave for all

three subjects. So, without actually estimating the parameters we may conclude that the model of Werner and Wooten cannot account for at least one kind of relevant data and should therefore be rejected.

Next we perform an analysis of the model of Larimer et al. (equation 5) along the same lines as we did with the model of Werner and Wooten.

First suppose the model generates a wavelength invariant unique green in the neighbourhood of 510 nm. This will be accomplished only if the contribution of the nonlinear part of the model is negligeable in that region and, since we take $x_1=-1$, if $x_2=\alpha(510)/\beta(510)>0$.

Secondly, in the long wave region where the contributions of the α- and β-systems are negligeable, ϕ_2 is positive so we must have $x_3>0$.

Following Larimer et al. we take $x_4<1$ in order to account for the wavelength shift of the short wave component in a nonspectral unique red.

With these restrictions in mind we now consider the capability of the model to generate concave iso-cancellation curves.

In terms of equation 5 an iso-cancellation setting corresponds to the implicit form

$$g(s,t) = ut + vs + \delta x_3 |(pt+qs)|^{x_4}$$

$$u = x_2\beta(\mu) - \alpha(\mu)$$

$$v = x_2\beta(\lambda) > 0$$

$$p = \gamma(\mu) - \beta(\mu)$$

$$q = \gamma(\lambda) - \beta(\lambda) > 0 \qquad (\lambda>550 \text{ nm})$$

of which the relevant partial derivatives are given by

$$g_s = v + \delta x_4 x_3 |(pt+qs)|^{x_4-1} q$$

$$g_t = u + \delta x_4 x_3 |(pt+qs)|^{x_4-1} p$$

$$g_{ss} = \delta x_4 (x_4-1) x_3 |(pt+qs)|^{x_4-2} q^2$$

$$g_{tt} = \delta x_4 (x_4-1) x_3 |(pt+qs)|^{x_4-2} p^2$$

$$g_{st} = \delta x_4 (x_4-1) x_3 |(pt+qs)|^{x_4-2} pq = g_{ts}$$

In order to evaluate the sign of f''(s) one has to evaluate the sign of these partial derivatives

Consider the case where $p \geq 0$. It is not difficult to see that in these cases $\delta=1$ and, since $q>0$, $g_{ss}<0$, $g_{st}<0$, $g_{tt}<0$, $g_t<0$ and $g_s>0$ hence $f''(s)<0$.

We were not able to think of any condition that would generate $f''(s) \geq 0$ so, for the time being, we accept equation 5 as a possible model for ϕ_2.

So far we have excluded two out of the three possible models for ϕ_2 In order to decide whether the remaining model of Larimer et al. (equation 5) is acceptable we have to consider its capability to account for the iso-cancellation curves quantitatively. We shall do so by estimating its parameters and seeing how well the model fits the data.

Parameter estimation

First, we shall discuss some general considerations concerning the way parameters of the nonlinear models should be estimated on the basis of cancellation data.

Generally, nonlinear models for cancellation data can be formulated as

$$F(s,t,x,\rho_\lambda,\rho_\mu) = 0$$

where x denotes the vector of parameters and ρ_λ and ρ_μ denote the vectors containing the output signals of the cone systems at wavelengths λ and μ respectively

In estimating x one may follow one of two possible strategies The first strategy is the most obvious and numerically the most simple It involves minimizing

$$\sum_i G\{F(s_i,t_i,x,\rho_\lambda,\rho_\mu)\} = \zeta$$

where G(y) is a function with exactly one minimum

This method has the disadvantage that it is difficult to judge the fit of the model because one minimizes not with respect to the data t_i but with respect to the transformation F of the data Clearly, the exact form of F depends upon the model

Furthermore, this strategy cannot be used if one wishes to compare different models on the basis of the actually obtained value ζ because different models involve different specifications of F and therefore values of ζ resulting from different F cannot be compared

In fact this is the strategy followed by Werner and Wooten and therefore their claim that the fit of their model is better than the fit of those of Larimer et al cannot be taken too seriously

The second strategy is numerically more complicated Let t_i^* be a real root of

$$F(s_i,t_i^*,x,\rho_\lambda,\rho_\mu) = 0,$$

then we minimize

$$\sum_{1} G(t_1, t_1^*) = \zeta^*$$

With t_1^* approximated by numerical methods (Szidarovsky and Yakowitz, 1978) equation (8) is an implicit function of the parameter vector x Normally, the derivatives with respect to these parameters can be obtained analytically by implicit differentiation so the minimization of ζ^* is amenable to standard iterative numerical methods (see also Note 3) (Jakoby, Kowalik and Pizzo, 1972) This strategy has the advantage that ζ^* can be used as a measure to compare the fit of different models applied to the same data

An more detailed appreciation of the fit can be obtained by comparing the values of t_1^* from the model with the optimal parameter vector x^* with the actual data t_1, for example by means of analysis of variance The resulting statistics will then reveal whether the data systematically differ from the predictions or not

For estimating the parameters of the model from the data reported in Experiment I, we used the second strategy and found that the resulting optimal parameter vector x^* was hardly, if at all, dependent upon the initial vector x^0

For $G(t_1, t_1^*)$ we took $(t_1 - t_1^*)^2/t_1$ because simple least squares would capitalize too much on large values of t_1

The results of the estimation procedure are shown in Table III From Table III it appears that the fit of the model of Larimer et al (equation 5) is not too good, the values of ζ^* are quite large, especially so for subject HE, and the squared correlations between measurements and predictions are

not high.

Table III

Summary of the results of estimating the parameters of the model of Larimer et al. (equation 5). r^2 denotes the squared linear correlation between data and the predictions associated with the optimal parameter set.

	CE	HE	CW
ζ^*	97.53	186.71	72.54
x_2	0.049	0.021	0.061
x_3	-0.044	-0.014	-0.059
x_4	0.983	0.978	0.979
r^2	0.71	0.21	0.74

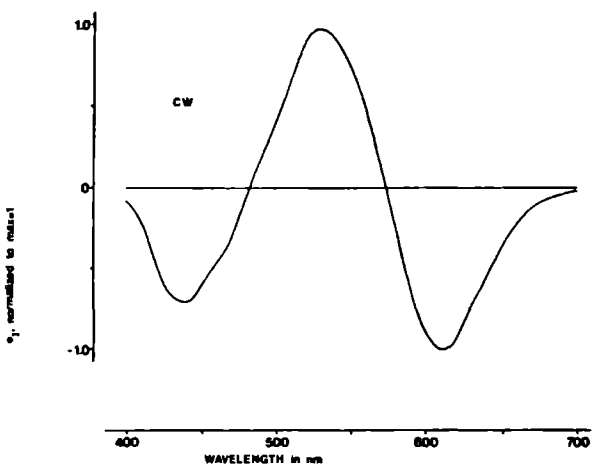


Figure 2. ϕ_2-curve for subject CW, simulated with the optimal parameter set for the model of Larimer et al. (equation 5) as obtained from the data of Experiment I. For explanation see text.

More serious is the fact that the estimates of x_3 are negative, implying that the optimal parameter set generates an extra equilibrium point for ϕ_2 in the long wave region. In Fig. 2 we plotted ϕ_2 as generated from the parameter estimates of subject CW. Clearly, the plot hardly resembles empirical ϕ_2-functions. Furthermore, according to the optimal parameter set the model is hardly nonlinear because the estimates of x_4 are very close to 1.0. From this analysis we conclude that although the model of Larimer et al. could be acceptable in a qualitative sense, it should be rejected because the optimal parameter estimates poorly fit the iso-cancellation data and leads to false predictions.

A new model

So far we have demonstrated that both models of Larimer et al. and the model of Werner and Wooten are not acceptable. We shall now discuss an other nonlinear model and show that it is qualitatively consistent with all relevant data. Furthermore, we will demonstrate that the model leads to acceptable estimates of the parameters and that it predicts an as yet unobserved phenomenon.

The new model is embodied in the equation

$$\phi_2(\lambda,s) = x_1\{s\alpha(\lambda)\}^{x_2} + x_3 s\beta(\lambda) + x_4 s\gamma(\lambda) \qquad (9)$$

with $x_2>1$.

Because beyond 500 nm $\alpha(\lambda)$ approximates zero the model generates wavelength invariance of unique green provided the parameters x_3 and x_4 are of opposite sign. The predicted wavelength of unique green will then depend upon the ratio $|x_4/x_3|$.

The model also predicts the observed wavelength shift of the short wave component as observed by Larimer et al. This can be observed from Fig. 3. Suppose that for a fixed radiance level we have a unique red setting, for example at μ=470 nm, and suppose we increase the total radiance while at the same time keeping the radiance ratio of the short and long wave components constant. Then the only way to compensate for the positively accelerated ($x_2>1$) activity of the α-system is to shift the short wave component in the direction of longer wavelengths.

To demonstrate that the model can also generate concave iso-cancellation functions we employ the same procedure as we did earlier.

Assume that the model indeed generates a wavelength invariant unique green and only positive values of ϕ_2 in the long wave region. Taking x_1=-1 this implies that

$$x_4>0, \quad x_3<0, \quad |x_3|=x_4$$

In terms of equation 9 an iso-cancellation setting corresponds to the implicit form

$$g(s,t) = -ut^{x_2} + sq + pt = 0$$

$$u = \alpha(\mu)^{x_2}$$

$$q = x_3\beta(\lambda) + x_4\gamma(\lambda)$$

$$p = x_3\beta(\mu) + x_4\gamma(\mu)$$

It is not difficult to derive that

$$f'(s) = -g_s/g_t = -q/\{-ux_2t^{x_2-1}+p\}$$

and that $g_{ss}=g_{st}=g_{ts}=0$ so

$$f''(s) = \{-g_s^2 g_{tt}\}/g_t^3$$

$$= \{q^2 u x_2(x_2-1)t^{x_2-1}\}/\{-ux_2 t^{x_2-1}+p\}^3$$

Now because $\lambda>550$ nm, $x_4>0$ and $|x_3|\approx x_4$ we have $q>0$ and since $f'(s)>0$ we derive $g_t<0$. Now $x_2>1$ is sufficient to conclude that $f''(s)<0$ so our model

Table IV
Summary of the results of estimating
the parameters of equation 9.

	CE	HE	CW
ζ^*	17.12	21.36	24.56
x_2	1.732	2.864	3.244
x_3	-0.019	-0.018	-0.016
x_4	0.013	0.013	0.011
r^2	0.965	0.942	0.0743
F_{1f}	12.52	13.84	19.72

will generate concave iso-cancellation curves and therefore is, at least qualitatively, superior to the models previously discussed.

We estimated the parameters of our model using the same method as we did before. The results are shown in Table IV. The most interesting thing to note about Table IV is that, although we did not constrain the parameter space, the properties of the parameter values found are exactly those we derived as sufficient for the model to be at least qualitatively acceptable,

i e

$x_2 > 1$ $\quad$ $x_4 > 0$

$x_3 < 0$ $\quad$ $|x_3| \approx x_4$

Also note that in terms of ζ^* and r^2 the fit of the model is better, be it only slightly for subject CW, than the fit of the model of Larimer et al

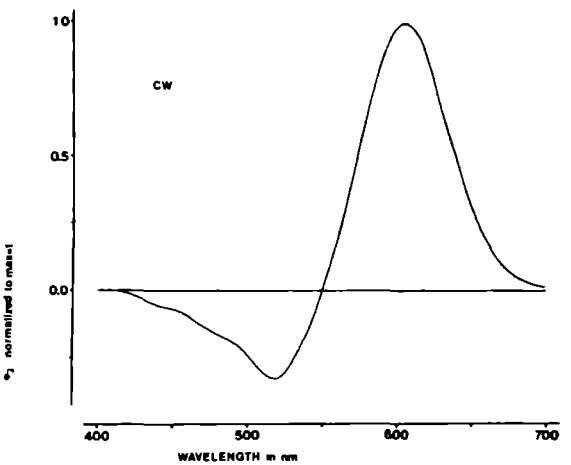


Figure 3 ϕ_2-curve for subject CW, simulated with the optimal parameter set for equation 9 as obtained from the data of Experiment I For explanation see text

In Fig 3 we plotted the ϕ_2-values for the entire spectrum as generated from the parameter values of subject CW Its general shape corresponds to the shape of empirical ϕ_2-functions although the equilibrium point is located at too long a wavelength We do not consider this to be a real flaw since the

wavelength of the equilibrium green was no constraint in the optimization of ζ^*. Implementing such a constraint did not affect the optimization results in any essential way We simulated the location of the equilibrium over a radiance range of 5 log-units and it appeared to be perfectly invariant The simulation results with the parameter set of subject CW are not essentially different from those obtained with the parameter sets of the other two subjects. On the basis of the qualitative analysis and the numerical results we conclude that, for the time being, our model is the only acceptable one

Further test of the model

From our model it can be seen that our model predicts that the wavelength of the short wave component of a unique red will shift towards *shorter* wavelengths if the total radiance increases and the radiance ratio of the short and long wave components is kept fixed and the wavelength of the short wave component is *shorter* than 445 nm. The reason is that if the total radiance increases the only way to compensate for the positively accelerated activity of the α-system is to either decrease or increase the wavelength of the short wave component. If this wavelength can only vary between 400 and 440 nm then the only possibility is to decrease the wavelength with increasing radiance

Note that none of the models that ascribe the nonlinearity to cone systems other than α would ever predict this kind of wavelength shift. Only those models that incorporate a *positive* accelerated activity of the α-system can predict wavelength shifts of the short wave component in two different directions.

In order to test this prediction we replicated the experiment of Larimer et al (1975)

Experiment II

Procedure

In this experiment we determined the wavelength of the short wave component in a mixture of variable short wave μ and a fixed long wave of 650 nm such that the resulting mixture neither appears bluish nor yellowish We determined such mixtures at three different radiance levels while over different radiance levels the radiance ratio of the short and the long wave was kept at a constant value of 4 0

Within each radiance level the subject was presented with a mixture of 650 nm and a variable short wave component with wavelength μ, ranging from 400 to 440 nm (this upper bound was chosen because it probably is lower than the wavelength of the peak sensitivity of the α-system) or from 460 to 500 nm 1 0 sec before stimulus onset the observer saw a LED-produced warning signal during 0 5 secs The stimulus remained visible during 4 0 secs At stimulus offset the observer had to indicate whether the stimulus appeared too bluish, too yellowish or that he was not able to decide between these two alternatives The interpresentation interval was 5 secs After each block of 5 presentations we presented the subject with a mixture that either looked too bluish or too yellowish in order to help him maintain a stable criterion The particular μ in a presentation was randomly selected on the basis of a staircase procedure (Cornsweet, 1962) We employed four different staircases, two of them starting at the lower bound (400 or 460 nm), the

other two at the upper bound (440 or 500 nm) As soon as the subject's judgement reversed from "too bluish" to "too yellowish" or vice versa, the direction of the staircase was reversed and the stepsize was halved On the other hand, as soon as the subject indicated that he could not decide between "too bluish" and "too yellowish" the stepsize was doubled The stepsize, however, never exceeded 5 nm and was never smaller than 1 25 nm Each staircase contained 15 steps This number was considered to be satisfactory because it appeared that with this number the stepsize in at least three out of four staircases was 1 25 nm at the last presentation Each presentation was randomly selected from one of the four staircases The wavelength μ that produces a unique red setting was estimated as the mean midpoint between successive opposite responses under the condition that the stepsize between those reversals was equal to or less than 2 5 nm

Observers

The same observers that participated in Experiment I also participated in this experiment except for subject CW, who was not available at the time of the experiment He was replaced by the color normal subject KS

Calibration

Since the density wedge W1 is not neutral (a gradual increase in density for wavelengths shorter than 450 nm) we obtained two-parameter transmission characteristics every 10 nm from 400 to 450 nm and interpolated the parameter values of the density characteristic for any wavelength between 400 and 450 nm by cubic splines (De Boor, 1978)

The mean wavelength settings per observer per radiance level are shown

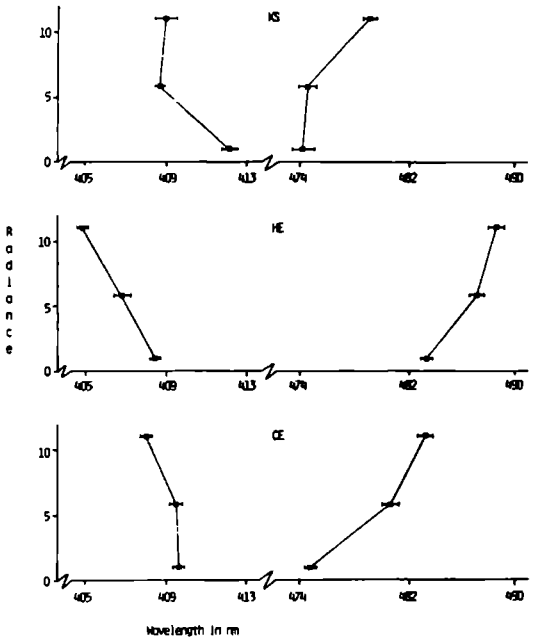


Figure 4. Results of experiment II. Wavelength μ of the short wave component of a unique red (long wave component at λ=650 nm) for different radiance levels of the short wave component with the radiance ratio of the short and long wave components constant (R(650)=4.0R(μ)). Radiance level 10 corresponds to 150.8 td.

in the wavelength-by-radiance plots of Fig. 4, together with the 80%-confidence intervals of those means. Clearly, with increasing radiance, the wavelength of the short wave component decreases when the variable short

wave component is shorter than 440 nm, just as predicted from our model. Furthermore, when the variable component is longer than 460 nm the results are consistent with those of Larimer et al. and those of Nagy. It is interesting to note that, although the optimization of ζ^* was not constrained

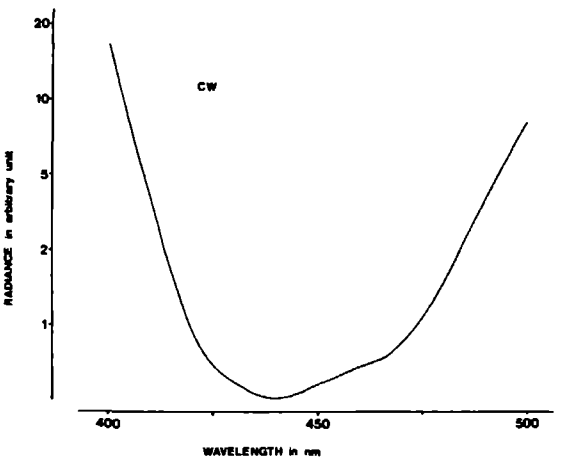


Figure 5. Simulated wavelength μ of the short wave component of a unique red with the long wave component at 650 nm and R(650)=4.0R(μ) on the basis of the optimal parameter set for equation 9 and the iso-cancellation data of subject CW. For explanation see text.

to generate these results, the parameter estimates allow for a qualitative reproduction of these results. We simulated the results of Experiment II by

taking x_2, x_3 and x_4 from Table IV and solving the equation

$$t^{x_2^*-1} = [x_3^*\{\beta(\mu)+4.0\beta(\lambda)\}+x_4^*\{\gamma(\mu)+4.0\gamma(\lambda)\}]/\{\alpha(\mu)+4.0\gamma(\lambda)\}^{x_2^*}$$

for each μ in the range from 400 to 500 nm and λ=650 nm. Of course, positive real solutions only exist if the right hand member of the above equation is not negative

The resulting solutions have been plotted for the parameter estimates of subject CW in Fig. 5 These results are also representative for subjects HE and CE Both the results of experiment II and the result of this simulation lend further credit to our model.

It is furthermore interesting to note that the simulation reported in this paragraph could not be performed with the parameter estimates of the model of Larimer et al since no real positive solutions exist for the appropriate equation in the relevant wavelength range

Discussion

A number of questions can be raised concerning the analysis of the various models treated in this paper

First, both for theoretical analysis and in the parameter estimation procedure we relied upon a set of hypothesized action spectra of cone systems for a standard observer Therefore, both the qualitative analysis of the models and the results of the parameter estimation procedure share all the possible flaws of these action spectra. However, the qualitative analysis of the models only relied upon a limited number of qualitative aspects of the Vos-Walraven primaries, features these primaries have in common with action spectra proposed by Smith and Pokorny (1975)

Furthermore, we used action spectra of a socalled standard observer although it is well known (Bastian, 1976; Alpern and Wake, 1977) that individual differences in the peak sensitivities of the cone systems may not be negligible and that interobserver differences exist in the density of the macular pigment and the ocular media. Although it is possible to determine directly the peak sensitivities of the cone systems (Alpern and Moeller, 1977) we felt that this procedure is too arduous to be practical. Without correcting for the precise location of these peaks it does not make much sense to correct for individual differences in the density of the eye media so we decided to use standard observer data only.

Second, the theoretical analysis of the various models partly relies upon wavelength invariance of unique green. The literature disagrees about this. Several authors present psychophysical (Dagher, Cruz and Plaza, 1958; Ingling and Tsou, 1977a; Nagy, 1979) data or physiological data (DeValois, Abramov and Jacobs, 1966) that indicate a wavelength shift of unique green towards shorter wavelengths with increasing radiance. On the other hand, Larimer et al. (1975) report wavelength shifts towards longer wavelengths for two color normals and wavelength invariance for two other color normals. Wavelength invariance is also confirmed by other data (Purdy, 1931; Moeller, 1976; Nagy, 1979). The results of Nagy and the analysis of Ingling and Tsou indicate that wavelength invariance only holds under conditions of prolonged exposure (0.5 secs. or more, see also Cohen, 1975) and photopic radiance levels. Thus we have the situation that our model might only, if at all, be valid for a limited set of viewing conditions: a wavelength shift of unique green towards shorter wavelengths cannot be explained by our model. Third, the fact that we ascribed the nonlinearity of the yellow/blue system to the

nonlinear contribution of the α-sytem might be due to the fact that we only considered the possibility of cone systems contributing to the opponent system It is quite possible that this apparent nonlinearity stems from a nonlinear contribution of the rod-system to the signal resulting from the α-system like suggested by Ingling and Tsou (1977a,b) However, it seems hard to explain the two-directional shift of the short wave component in a unique red setting in terms of a rod signal intrusion to the α -system in the sense of Ingling and Tsou.

Of course, whether or not rod activity does play a part in the activity of the yellow/blue mechanism this does not imply that the only way to account for our results is a model like we propose It is very well possible that the contribution of the β- and ϒ-cone systems is nonlinear also A more complete understanding of the mechanism could be obtained by studying additive closure in a systematic way by means of the paradigm used by Ikeda and Ayama (1980). Our results only indicate that without a nonlinear contribution of the α -system some data cannot be understood.

An even more fundamental question is whether it is still worthwhile to pursue the explanation of visual experience in terms of an opponent process theory that employs the concepts of "redness/greenness" and "yellowness/blueness" as its building blocks since it appears that these building blocks themselves do not all possess a well defined sensitivity function, i e the sensitivity function for ϕ_2 does not satisfy

$$\phi_2(\lambda,t) = t\phi_2(\lambda,t)$$

for all real values of t and all wavelengths λ Therefore (Elzinga, 1984) models for ϕ_2 cannot be dimensionally invariant and the ϕ_2-function for a

standard observer as proposed by Werner and Wooten (1979b) is of no use at all.

The analysis of variance employed in evaluating the adequacy of different models is based upon the following considerations.

Each measurement t_{ij} $(i=1,..,k;j=1,..,n)$ can be conceived of as a linear combination

$$t_{ij} = M(s_i) + \omega_i + \varepsilon_{ij}$$

where $M(s_i)$ now represents some specified model, ω_i represents the bias of $M(s_i)$ and ε_{ij} random error. If the specification of M is correct, each $\omega_i=0$, if not, the value of ω_i depends upon M.

Provided the variance of the t_{ij} per level s_i is a constant the quantity

$$MS_W = \sum_{ij} (t_{ij}-t_{i.})^2/k(n-1)$$

is an unbiased estimate of σ_ε^2.

If the model is correct the quantity

$$MS_M = \sum_{i} \{(t_{i.}-M(s_i)\}^2/(k-p)$$

where p is the number of parameters of M, provides an unbiased estimate of σ_ε^2 also. If, however, the model is not correct the expectation of MS_M equals

$$\sigma_\varepsilon^2 + \sum_{i} \omega_i^2/(k-p)$$

So, if the ε_{ij} are normally distributed, homoscedastic and independent the ratio

$$F_{lf} = MS_M/MS_W$$

represents a sample from an F-distribution with k-p and k(n-1) degrees of freedom.

Note 2

Werner and Wooten (1979a,b) do not mention the sign function δ in their model. However they must have used it because without that sign function ϕ_2 would have two local maxima beyond 500 nm and a deep trough around the point where $x_2\beta(\lambda)=x_3\gamma(\lambda)$. Such a trough is absent both in their data and in the plots of ϕ_2 that they generate with their model.

Note 3

Since the parameters of nonlinear models have to be estimated by an iterative algorithm, the algorithm has to be provided with an initial guess of x: x^0. In order to evaluate ζ^* at x^0 or, in any subsequent iteration j, at x^j the equation $F(s_i,t_i^j,x^j,\rho_\lambda,\rho_\mu)=0$ has to be solved for a real root t_i^j by means of a second (inner) iterative process. However, such a real root may not exist for some x^j and or all values of s_i, depending upon the nature of the model.

In fact this phenomenon occurred with our model. We circumvented the problem by solving for the real roots of

$$-\delta|t_{ij}^j\alpha(\mu)+s_i\alpha(\lambda)|^{x_2}+t_{ij}^j\{x_3^j\beta(\mu)+x_4^j\gamma(\mu)\}$$
$$+s_i\{x_3^j\beta(\lambda)+x_4^j\gamma(\lambda)\}=0$$

where j denotes the outer iteration phase and $\delta=\text{sign}\{t_i^j\alpha(\mu)+s_i\alpha(\lambda)\}$. In case $\delta=-1$ the solution t_i^j is always negative so the contribution of these solutions to the value of ζ^* is relatively large which in turn forces the

algorithm to move towards parameter vectors $x^{\{j+i\}}$ that allow for real solutions with $\delta=1$. This method has the effect of introducing a penalty function into the outer algorithm without introducing discontinuities in the model itself or into the derivatives with respect to x or the roots t^j.

Note further that the above equation may have, depending on the values of x^j and s_i, one, two or three real roots. Whenever the equation has more than one real root we chose that root for which $|t_i^j - t_i|$ is minimal. As a result of this the procedure as a whole is rather slow.

At the time the present chapter was written, the evidence on the nonlinearity of the yellow/blue code, as presented in Chapters 2 and 4, was not available. The analysis of the experimental results and the theory presented in Chapter 5 explicitly assumes that the hue-cancellation function ϕ_2 for the yellow/blue system is linear.

However, the assumption that the yellow/blue system is nonlinear and behaves according to the model presented in Chapter 4 (equation 9), does not alter the analysis or conclusions of the present chapter in any significant way since the nonlinearity of the yellow/blue system only becomes apparent with stimuli that contain a wavelength component shorter then, approximately, 510 nm. Beyond that wavelength the yellow/blue code behaves linear since the activity of the short-wave sensitive cone system is negligible in that part of the spectrum.

We, at that time, furthermore assumed that flickerphotometric luminance is a code that is linear with respect to color-mixture operations. Work done by Ingling, Tsou, Gast, Burns, Emerick and Riesenburg has shown that this assumption is untenable but that the deviations from linearity are very small. Considering the fact that the variation in magnitude estimation data is tremendously large, compared with these minor deviations, we think that taking these deviations into account would not lead to the slightest change in any of the conclusions drawn in this chapter.

HUE MAGNITUDE ESTIMATES AS RELATIVE JUDGEMENTS[1]

Abstract

A simple model in terms of nonlinear opponent color codes for the veiling of a weak hue component by a strong one in mixtures of unique yellow and unique green is explored experimentally by means of a procedure which is an extension of the classical hue magnitude estimation procedures. The model is extended towards a more general model for the veiling phenomenon. An experimental validation of the procedure is included. Independent evidence for a hypothesis of Guth and Lodge (1973) regarding inhibition of the red/green opponent system by the yellow/blue opponent system was found when brightness estimates were conceived of as a function of the log luminance of the yellow stimulus component.

Introduction

Opponent colors theories assume that the linear nonopponent output signals of three photopigments are recoded into three linear signals, namely a whiteness/blackness signal, a red/green signal, and a yellow/blue signal. To account for the Bezold-Brücke hue shift and Abney's hue shift, it is furthermore assumed that these opponent linear signals are recoded into at least three nonlinear signals. For an excellent substantial and historical review of opponent colors theory, the reader is referred to Hurvich (1977). Krantz (1975a, 1975b) first provided a formal basis for opponent colors theory which is at the same time a measurement theory for the cancellation-energy measurements of Hurvich and Jameson (1957; Jameson & Hurvich, 1955).

---

[1]Elzinga, C.H. & De Weert, Ch.M.M. (1978). Hue magnitude estimates as relative judgements. *Perception & Psychophysics, 23 (5),* 372-380.

Within the set of all lights, A, Krantz distinguishes two subsets, A_1 and A_2. The set A_1 consists of all lights that are neither reddish nor greennish; A_2 consists of all lights that are neither bluish nor yellowish. Consequently, the set $A_1 \cap A_2$ consists of all lights that appear colorless. All lights $a \epsilon A_1 \cap A_2$ are called equilibrium lights or unique colors. If $a \epsilon A_1$, a is called unique yellow or blue; if $a \epsilon A_2$ a is called a unique red or green.

If the axioms of Krantz' theory are satisfied, A can be represented by a convex cone, C in a linear three-dimensional space, i.e., there exists a homomorphism, ϕ, which maps each light $a \epsilon A$ onto a vector $\phi(a)=\{\phi_1(a),\phi_2(a),\phi_3(a)\} \epsilon C$. The functions (coordinates of C) ϕ_1 and ϕ_2 represent the linear form of the red/green and yellow/blue opponent cancellation codes; i.e., the functions $\phi_i (i=1,2)$ are linear with respect to color mixture operations:

$$\phi_i\{(u*a)\odot(v*b)\}=u\phi_i(a)+v\phi_i(b) \qquad (1)$$

for all $a,b \epsilon A$ and all $u,v \epsilon Re^+$. Here * and $\odot$ denote, respectively, an overall change in energy level and additive mixture of colored lights. In fact, the connotation of Equation 1 is that ϕ maps the cone $\langle A,\odot,* \rangle$ onto the cone $\langle Re^+,+,\times \rangle$.

An important property of the ϕ_i is that $\phi_i(a)=0 \leftrightarrow a \epsilon A_i$ for i=1,2. From the viewpoint of empirical validity of Krantz' axiomatization, of interest are those axioms which imply closure of the A_i under the operations $\odot$ and *. These axioms have been tested extensively, under dark adaptation, by Larimer, Krantz, and Cicerone (1974, 1975). Krantz' axioms are independent from the state of adaptation of the eye in the sense that closure of the A_i

should hold for any state of adaptation. However, the make-up of the A_i may, and in fact does, change with changes of the state of adaptation. Cicerone, Krantz, and Larimer (1975) tested for closure of the A_i under different conditions of color adaptation.

Krantz' axioms are a set of axioms which describe a state of affairs in a physical space; now the psychophysical problem is to establish the relation between the opponent linear cancellation codes, ϕ_i, and the apparently nonlinear codes, f_i, which occur later on in the visual system and which are related to subjective color experience. This problem could also be formulated as the problem of establishing functions, F_i, that map the ϕ_i onto codes for subjective color experience, f_i. These transducer functions, F_i, cannot be one-argument functions because of what Krantz has called the veiling phenomenon.

Veiling is the masking of a weak hue component by a strong one. Imagine a unique green light, a, and a unique yellow, b. According to Krantz' axioms, the red/green cancellation code for b must equal zero because $b \epsilon A_1$: $\phi_1(b)=0$; and because $a \epsilon A_2$, we also have $\phi_2(a)=0$. But because both lights are colored, we must have $\phi_2(b)>0$. Now consider the mixture $a \odot b$: because ϕ_1 must satisfy Equation 1 we have, with $\phi_1(b)=0$, $\phi_1(a \odot b)=\phi_1(a)+\phi_1(b)=\phi_1(a)$.

Thus the red/green cancellation code ϕ_1 has the same numerical value for a and the additive mixture $(a \odot b)$. But direct magnitude estimation of the greenness of a and $(a \odot b)$ would reveal that the subjective impression of greenness evoked by $(a \odot b)$ is considerably less than that evoked by a alone. Raising the physical intensity of b in the mixture would not alter the red/green cancellation value of the mixture, but would cause the subjective impression of greenness from that mixture to vanish almost completely. Thus

the addition of yellow to green seems to veil the subjective impression of greenness, and thus the veiling phenomenon seems to rule out any expression of the form $f_i=F_i(\phi_i)$. The existence of a veiling phenomenon was already suggested by Yager and Taylor (1970) when their empirical hue coefficient function grew faster than was predicted on the basis of their estimated luminance thresholds for perceived yellowness and greenness and the estimates (HMEs) were expressed as power functions of the luminance. However, their estimates of these exponents were about four times smaller than the exponents estimated by Raaijmakers and De Weert (1975). The latter conducted an experiment in which they obtained HMEs of greenness for different values of the luminance of a unique green stimulus under admixture of two levels of a unique yellow. The HMEs clearly seemed to show veiling of the green by the admixed yellow. However, by describing their data as a function of the proportion of the luminance of the green in the mixture, it seemed as if no veiling had occurred at all. They concluded that their subjects probably had not been able to give absolute judgments about the amounts of green in the stimuli. Raaijmakers and De Weert suggested that the relative strength of the chromatic and achromatic signals might be the only measures available to the subject, so that the HMEs produced by their subjects could not be regarded as absolute judgment about the nonlinear opponent code for green. A further indication that this might be true is the fact that the logarithm of the power functions as obtained by Yager and Taylor grew faster at lower luminance levels and only became linear with log-luminance at higher luminance levels. If the suggestion of Raaijmakers and De Weert, that only information about the relative strenght of chromatic and achromatic signals is available to the subject, a rather simple model

for hue magnitude estimates in terms of the nonlinear opponent color codes f_i is given by

$$H_i = cBf_1/(f_1+f_2) \qquad (2)$$

Here H_i denotes the hue magitude estimate associated with f_i, and cB is a positive scalar transformation of the brightness estimate, B. We shall assume that the f_i are strictly monotone with luminance and thus with the ϕ_i. With this very weak assumption, it is easy to see that this model recognizes veiling: if, given a constant ϕ_1 level of a mixture, ϕ_2 increases, f_2 will also increase, thus diminishing the ratio $f_1/(f_1+f_2)$, and, given a suitable choice of c, H_1 will decrease. This corresponds to the veiling of greenness caused by raising the intensity of a unique green. For a constant ϕ_2 level, H_1 will be monotone with ϕ_1.

Notice that equation 2 assigns special status to brightness as a psychologically more primitive aspect of light than hue. However, this is not meant to say that the brightness sensation arises independently from the f_i. If we denote the nonlinear opponent code for whiteness/blackness as f_3, brightness could be thought of as a function of all three nonlinear opponent codes: $B=h(f_1,f_2,f_3)$. Such a proposal has already been made by Guth, Donley, and Marocco (1969).

The experiment to be reported here concerns the evaluation of equation 2 as a candidate for explaining the veiling phenomenon without accepting the assumption that the nonlinear opponent hue codes are multiargument functions of the linear opponent codes.

Method

Apparatus

In Figure 1, a diagram of the equipment is given

Two xenon arc lamps were used as light sources. The light of these sources was transformed into monochromatic light with the help of two Carl-Leiss prism monochromators. Monochromator M1 provided the yellow beam,

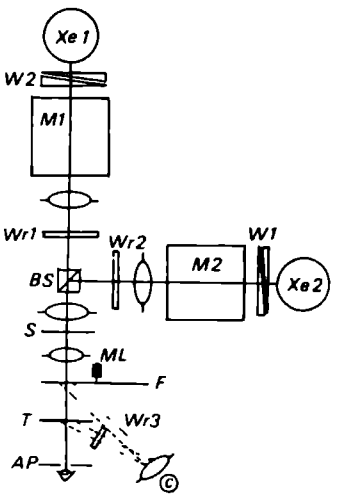


Figure 1. Diagram of the optical equipment. Monochromators M1 and M2 are fed by 150-W xenon arc lamps Xe1 and Xe2. Wr1 and Wr2 are neutral density filters; W1 and W2 compensated circular neutral density wedges, BS, beamsplitter; S, electromagnetically driven shutter, F, flickervane, T, test target, subtending 1.5°; Wr3, neutral density filter, AP, 2-mm artificial pupil; C, calibration light source.

monochromator M2 the green one. The intensity of the sources was controlled by varying circular neutral density wedges W1 and W2. Through beamsplitter BS, the yellow and the green beam were mixed. With the help of lens ML, the target was seen in Maxwellian view and subtended 1.5°. The various mixtures of the luminances of the yellow and the green beam were selected by choosing the appropriate neutral density filters from Wr1 and Wr2. Wr1 and Wr2 both consisted of a disk in which a number of different density filters were placed. By rotating the disks, different filter combinations were chosen. S is an electromagnetically driven shutter with a circular diameter of 5 cm. A source, C, could be projected onto a flickervane, F, for the purpose of equating the yellow and the green beam for luminance. C was also used for a constant and faint illumination of the surroundings of the target. The brightness level of the target surroundings was kept constant during the whole experiment and was much lower than that of the least bright stimulus.

Subjects

Three subjects, A, C, and F, served in the experiment. None of them had any prior experience with color experiments. All subjects had normal color vision. The subject used his right eye; the left eye was kept covered during the whole experimental session. The head of the subject was fixed by using a chin- and foreheadrest.

Determination of Unique Colors

After a dark-adaptation period of 10 min, a rough determination of the wavelengths of the unique colors was made. Following a readaptation period of 3 min, seven different wavelengths around the estimated unique

wavelengths were presented in a random-block design. Each wavelength was presented 10 times.

Table 1

Wavelength settings for unique yellow and unique green

Subject	Yellow (nm)	Green (nm)
A	586	513
C	581	520
F	574	516

The subjects had to say whether red or green was present in the case of yellowish stimuli and, in the case of greennish stimuli, whether yellow or blue was seen. The 50%-50% value was taken as the wavelength of the unique color. In Table 1, the wavelength values, at a retinal illumination level of 168 photopic trolands, of the unique colors of each subject are presented.

Procedure

After determination of the wavelength of the unique colors as described above, the luminances of both colors were equated according to the flickerphotometric procedure at the 0.0 values of the density filters. Subsequently, the subject was instructed about his task as follows: "Each 30 seconds, after a click is heard, a short flash will be given, the color of which varies between yellow and green." The range of colors was then shown to the subject. "Your task is to assign a number to the brightness of these flashes and to assign numbers to the amount of green and the amount of

yellow present in the flash. These judgments should feel right. You may use whole numbers, fractions, or a combination of the two. Try to make the numbers directly proportional to your impression of the brightness, the amount of green and the amount of yellow present in the flash. So, for each flash, you have to judge three different aspects of the flash: the brightness, the amount of yellow, and the amount of green. Assigning a number to the brightness, you should try to ignore the greenness and the yellowness of the flash; in assigning a number to the amount of green (yellow), you should try to ignore the brightness and the amount of yellow (green) in the flash. Constantly try to remember that the number you use should be proportional to your own subjective impression of the brightness and the amounts of each hue present in the flash. All flashes are mixtures of yellow and green, so you should avoid the use of the number zero. Try not to worry about being consistent, small variations are quite normal for this kind of task." Thereafter, the subject was asked to restate the instruction in order to test his comprehension of the task.

The stimuli were presented in random order, and the order of the judgments per presentation was also randomized. Each stimulus was presented five times. Stimuli were presented for 4 sec. Five hundred milliseconds before presentation, one 20-msec click was presented as a warning signal.

From the first 15 presentations, a single stimulus was randomly selected. This stimulus was presented after each 15th presentation and the subject was then informed about the brightness and hue magnitude estimates he had produced at the first presentation of this stimulus.

Range of yellow and green luminance

In Table 2, six different levels of luminances of the yellow and the green are presented. The 36 combinations of these levels were randomly presented, each combination five times.

Table 2

Luminance values of the stimulus components (in Trolands)

Green	1.5	3.5	8.5	21	53	168
Yellow	1.5	3.5	8.5	21	53	168

Validation of the Experimental Method

In fact, our experiment consisted of three different experiments as three different types of judgments were required from the subjects. One would expect that, if this procedure indeed corresponds to three separate experiments, hue magnitude estimates and brightness estimates from separate experiments, should be scalar transformations of the corresponding estimates, as obtained in the combined experiments. Subject F was used to test this hypothesis. Subject F served in three separate experiments which were an exact replication of the experiment described above, except for the fact that in each of these experiments only one of the three kinds of estimates was required from the subject. The replications were run on 3 separated days in succession and 4 days after the original experiment with subject F had taken place. To evaluate our prediction that the estimates from the original experiment were scalar transformations of the brightness

and hue magnitude estimates obtained from the replications, we tried to predict the results of the combined experiments from the results of the replications according to the model $Y=\beta X+\varepsilon$. As a measure of the goodness of fit of this model, we use a coefficient of determination[1] (d) which is a direct analogue of the squared correlation coefficient (r^2). d was used instead of r^2 because what matters is proportionality between X and Y and not primarily linearity. In Table 3, the coefficients of determination for the three types of judgments, together with the squared correlation coefficients and the estimated scalars (β) are presented.

Table 3

Values of the coefficient of determination (d), the squared correlation coefficients (r^2), and the estimated scalars (β) resulting from the prediction of magnitude estimates obtained from the combined experiments out of magnitude estimates of three separate experiments.

Judgment	d	r^2	beta
Brightness	.82	.90	1.19
Amount of Yellow	.73	.74	.82
Amount of Green	.80	.83	1.05

As can be seen from Table 3, the fit of the data to the hypotheses of linear transformation is slightly better. The better fit of a linear model might be explained by assuming time-dependent response criteria. The figures in Table 3 seem to justify the conclusion that our method leads to the same kind of judgments as those which would have been obtained by running a separate experiment for each type of judgment. If this conclusion is correct, indeed, then our procedure has led to a considerable decrease in

experimental effort as compared to conventional procedures of magnitude estimation in which only one type of judgment is required from the subjects.

Theory of Conjoint Measurement

The theory of conjoint measurement concerns the way in which independent variables determine a joint effect according to one or another rule of composition of independent variables. Here we shall be concerned only with those aspects of the theory which are directly relevant to the subject matter of this article. The interested reader is referred to Krantz, Luce, Suppes, and Tversky (1971) for a detailed and complete treatment of conjoint measurement structures.

Let G and Y denote, respectively, the set of all unique green lights and the set of all yellow lights, and let M denote the Cartesian product matrix of $G \times Y$, of which the elements $\{m(a,w)\}$ are numbers representing the joint effect of the additive mixture of $a \in G$ and $w \in Y$, i.e., $G \times Y$ is weakly ordered and m is an ordinal scale on $G \times Y$.

M satisfies monotonicity (or independence) whenever, for $a,b \in G$ and some $w \in Y$, $m(a,w) \geq m(b,w)$ implies $m(a,x) \geq m(b,x)$ for all $x \in Y$ and, for $x,y \in Y$ and some $c \in G$, $m(c,x) \geq m(c,y)$ implies $m(d,x) \geq m(d,y)$ for all $d \in G$. If M is an $n \times n$ matrix, there are $2\binom{n}{2}(n-1)$ possible tests of monotonicity. For all practical purposes, the above mentioned definition implies that a matrix M satisfies monotonicity if there exists a permutation of rows and columns of M such that within each row the elements are nondecreasing from the left to the right and such that, within each column, the elements are nondecreasing from top to bottom.

M satifies double cancellation whenever, for all $a,b,c \in G$ and all $w,x,y \in Y$,

$m(a,x) \geq m(b,w)$ and $m(b,y) \geq m(c,x)$ together imply $m(a,y) \geq m(c,w)$. If M is an $n \times n$ matrix there are $\binom{n}{3}^2$ triples to be checked with respect to double cancellation. These tests are not independent, since the conclusion of one may be one of the antecedent of another. For some of the triples, a test for double cancellation may not be possible because, for some triples, the antecedents may not allow for the conclusion of the double cancellation rule. If M does not satisfy double cancellation, the violations may be one of two possible kinds of violations. A "strong" rejection occurs if both antecedents are strict inequalities and the conclusion doesn't hold. If, however, one of the antecedents is an equality and the conclusion doesn't hold, there is a violation also. This type of rejection is called "weak" , because with finer grained data the equality antecedent might go either way, and in one of these cases the antecedents will not be apt to allow for a test of double cancellation.

If M satisfies, apart from a number of axioms of which the empirical consequences are hard to evaluate, monotonicity and double cancellation, it can be proven, as shown by Krantz et al. (1971), that real-valued functions ψ_1, ψ_2 exist such that $m(a,w) = \psi_1(a) + \psi_2(w)$. The functions ψ_1 and ψ_2 are unique up to positive linear transformations with a common unit. There are several, rather laborious, methods for the construction of standard sequences that provide solutions for ψ_1 and ψ_2. For a discussion of some of these methods and the application of one of them, the reader is referred to Levelt, Riemersma, and Bunt (1972). Alternatively, a suitable nonmetric multidimensional scaling algorithm may be employed to solve for the scales ψ_1 and ψ_2.

The most direct test of equation 2 is to test the hypotheses that $H_i+H_j = cB$ (c denotes a positive constant). The results of this test appear in Table 4 and Figure 2.

Table 4

Results of testing $H_i+H_j=cB$

Coefficients of determination, squared correlation coefficients and estimated scalars resulting from the prediction of brightness estimates from the sum of the hue magnitude estimates.

subject	d	r^2	c
A	.87	.93	.89
C	.95	.95	.90

In Table 4 the d's refer to the prediction of the brightness estimates from the sum of the HMEs. The lines in Figure 2 represent the predicted brightness estimates. Clearly, these data support our hypotheses. Equation 2 also leads to

$$\log(H_i/H_j) = \psi_i(\log f_i) - \psi_j(\log f_j), \qquad (3)$$

which can be considered as a subtractive conjoint measurement model. If a subtractive representation for $\log(H_i/H_j)$ exists, the matrix of ranks of $\log(H_i/H_j)$ should at least satisfy monotonicity and the double cancellation rule. These properties were tested for both subjects A and C. The results of these tests appear in Table 5. As can be seen from Table 5, the matrices of ranks of $\log(H_i/H_j)$ of both subjects do not satisfy monotonocity, nor does the matrix of subject A satisfy double cancellation.

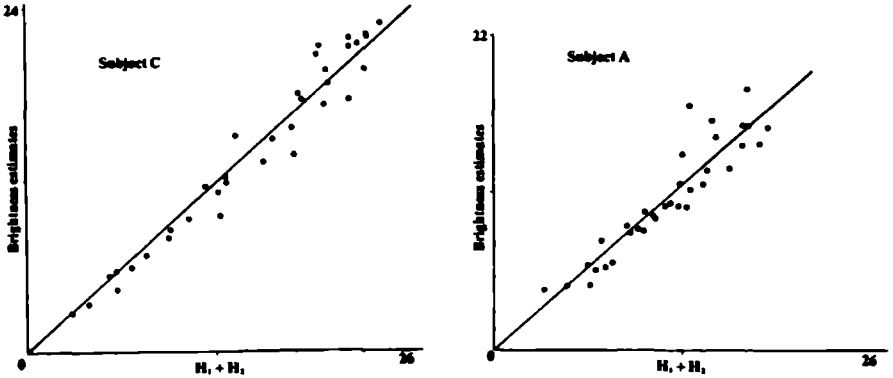


Figure 2. Brightness estimates vs. the sum of hue magnitude estimates. Points represent the geometrical means of five presentations.

Table 5

Results of the tests for monotonicity and double cancellation of the matrices of ranks of $\log(H_i/H_j)$ for the complete set of stimuli.

	6 by 6 Matrix	
Cancellation Test	A	C
Acceptance	399	400
Weak Rejection	0	0
Strong Rejection	1	0
No Test Possible	0	0
Number of violations of monotonicity	11	9

Note: Total number of attempts of double-cancellation tests is $\binom{6}{2}^2=400$. Total number of tests of monotonicity is $\binom{6}{2}10=150$.

However, in almost all the violations, the first row and column of these matrices are involved No permutation of rows and columns leads to a substantial improvement of monotonicity without at the same time producing an increase in the number of violations of double cancellation. As the first row and column in these matrices refer to stimuli in which the physical amounts of yellow and/or green were highest, the violations of the necessary conditions of monotonicity and double cancellation could be considered as a response bias, caused by a restriction of the response range It could also be true that equation 2 is simply false for higher luminance levels. In any case, the next question is whether a more restricted set of stimuli could lead to a subtractive representation. We, therefore, computed the matrices of ranks of $\log(H_i/H_j)$, ignoring those stimuli in which the physical amounts of yellow and/or green were maximal From the figures in Table 6, it is clear that these matrices turn out to satisfy monotonicity (almost) perfectly

Table 6

Results of the tests for monotonicity and double cancellation of the matrices of ranks of $\log(H_i/H_j)$ for the reduced set of stimuli

	5 by 5 Matrix	
Cancellation Test	A	C
Acceptance	100	100
Weak Rejection	0	0
Strong Rejection	0	0
No Test Possible	0	0
Number of violations of monotonicity	4	0

Note-Total number of attempts of double-cancellation tests is $\binom{5}{2}^2=100$ Total number of tests of monotonicity is $\binom{5}{2}8=80$

Now the conclusion seems justified that an acceptable solution exists for a subtractive representation for the reduced set of stimuli. Such a solution was obtained by using an algorithm which was developed by Roskam (1974) and can be considered as an extension of nonmetric multidimensional scaling algorithms such as Kruskal's.

If a subtractive representation for $\log(H_i/H_j)$ exists, scales $\psi\{\log(f_i)\}$ can be found which are unique up to positive linear transformations with a common unit. Conjoint analysis of the ranks of $\log(H_i/H_j)$ thus produces scales for f_1 and f_2 of the form

$$\log f_i = \{\psi_i(\log f_i) - \sigma_i\}/\alpha.$$

As we are interested in the f_i, the additive constants σ_i have to be estimated (α can arbitrarily be chosen to equal 1). According to equation 2, it must be true that $H_i/H_j = f_i/f_j$, so the parameters σ_i and σ_j can be estimated by minimizing the function ξ:

$$\xi = \sum_i \sum_j \{\log(H_i/H_j) - (\psi_i - \sigma_i - \psi_j + \sigma_j)\}^2.$$

The least squares estimates of the σ_i were obtained by using an algorithm developed by James and Roos (1972).

Having obtained these parameter estimates, estimates of the $\log f_i$ are easily obtained by subtracting the estimated constants from the corresponding scales ψ_i. In Figure 3, these estimates are plotted against the logarithm of the luminance of the stimulus components by L_i and assuming that

$$f_i = k_i L_i^{\beta_i},$$

the coefficients β_i can be estimated by simple linear regression.

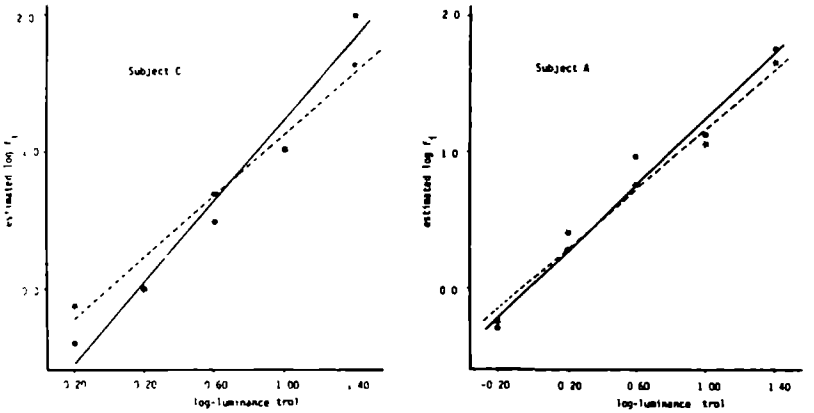


Figure 3. Estimated $\log(f_1)$ values vs. the log luminance of the stimulus components. The points represent the estimated $\log(f_1)$-values, the stars represent the estimated $\log(f_2)$-values

These estimates are presented in Table 7 for both subjects, together with the squared correlation coefficients.

Table 7

Values of the estimated exponents (β_i) when opponent nonlinear codes f_i are conceived of as power functions of the luminance of the stimulus components.

Subject	beta1	r21	beta2	r22
A	1.12	.98	1.23	.97
C	1.13	.98	1.48	.98

Furthermore, if equation 2 is correct, the HME's should be predictable from the estimated f scales and the observed brightness estimates by

$$H_i = cB\hat{f}_i/(\hat{f}_i + \hat{f}_j)$$

Here f_i denote the estimated opponent nonlinear codes. In Figure 4 plots of the predicted HMEs against the observed HMEs are given for both subjects. In Table 8, the coefficients of determination and the squared correlation coefficients associated with these predictions are presented.

The fact that these squared correlation coefficients are somewhat higher than the coefficients of determination might be explained by assuming that, at the lowest luminance levels, subjects restrict the lower bound of their response range to the number 1 and are not willing or able to use fractions smaller than one. The plots in Figure 4 do suggest such a response effect.

Table 8

Coefficients of determination (d) and squared correlation coefficients (r^2) resulting from the predictionof the observed hue magnitude estimates from the estimated scales for the opponent nonlinear codes

	Yellow		Green	
Subject	d	r^2	d	r^2
A	.92	.98	.95	.98
C	.94	.98	.96	.98

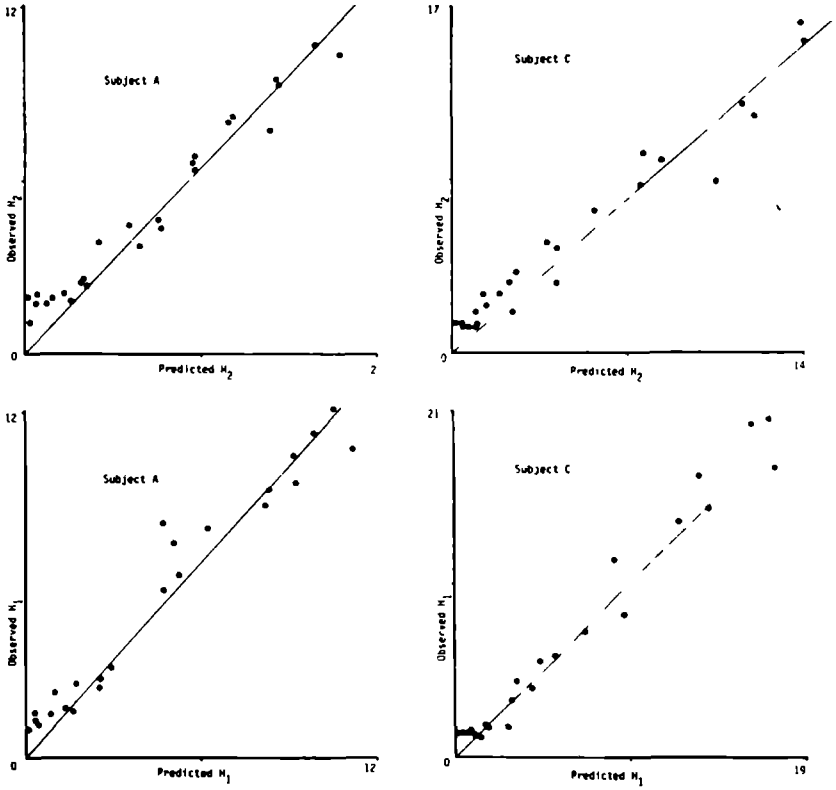


Figure 4. Observed magnitude estimates vs magnitude estimates as predicted from the estimated f_i values and the brightness estimates

On the basis of the aforementioned results, it seems safe to conclude that our model for hue magnitude estimates has not been refuted by our data. This indicates that the veiling phenomenon should be considered as the result of the response generating process and not of the signal transducing mechanism. Interesting as this may be, the model, as such, in no way can be considered as a complete account of the veiling phenomenon, because, as reported by Hurvich and Jameson (1951), broad-band light is judged white (colorless) if the luminance is raised sufficiently. This phenomenon should be interpreted as veiling of a hue component by the whiteness/blackness code, f_3. A rather direct extension of our model which is capable of describing hue veiling by whiteness is given by

$$H_i = cB(f_i / \sum_j^3 w_j f_j)$$

In equation 4, the w_i are scalars which determine the relative contribution of each f_i to the denominator of the equation. That our model still works quite well for the data from our experiment indicates that w_1 and w_2 are approximately equal and large as compared to w_3. Equation 4 directly corresponds to an experiment that is analogous to the one described above. However, even equation 4 cannot be considered as a complete model in the sense that we do not have a theory for the brightness estimates in terms of the f_i, i.e. we cannot specify h in $B = h(f_1, f_2, f_3)$. Although a specification of h in terms of the f_i does not seem to be possible at this moment, some attempts have been made to specify B in terms of the cancellation energies, notably by Guth and Lodge (1973). In this same article Guth and Lodge speculate about the possibility of inhibition of the

red/green system by the

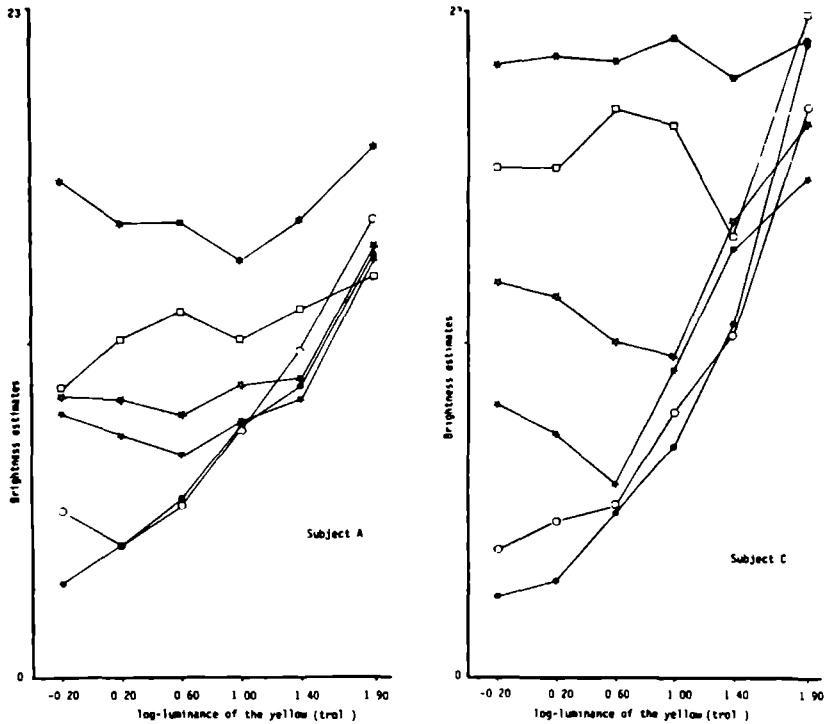


Figure 5. Brightness estimates vs. log luminance of the yellow stimulus component for different levels of the luminance of the green component (L_{gr}). ✱, L_{gr}=168 trolands; □, L_{gr}=53 trolands, ✲, L_{gr}=21 trolands; ✲, L_{gr}=8.5 trolands; ○, L_{gr}=3.5 trolands; ●, L_{gr}=1.5 trolands. Each point represents the mean of five presentations.

yellow/blue system. Our data suggest the existence of such an effect. In

Figure 5, the brightness estimates are plotted against the logarithm of the luminance of the yellow component in the stimuli. A simple two-way analysis of variance of the brightness estimates reveals a significant interaction effect of the yellowness and blueness on brightness. From Figure 5, it seems as if the brightness responses are inhibited by the luminance of the yellow as long as this luminance is lower than the luminance of the green component. This indicates that the yellow/blue system inhibits the red/green system for a restricted range of the relative luminances of the yellow and the green. If the hypothesis of red/green inhibition by the yellow/blue system is accepted, it should be incorporated in a model for brightness estimates in terms of the nonlinear opponent color codes, f_1. In order to arrive at a more complete understanding of the veiling phenomenon, it might be interesting to try to quantify veiling itself for example, by measuring the energy that has to be added to a unique green in a mixture of a unique green and a unique yellow in order to make the mixture contain an equal amount of greenness as the unmixed green did.

1. For the model $Y=\beta X+\varepsilon$, a least squares estimate of β is provided by $(\Sigma XY)/(\Sigma X^2)$, which implies as an estimate of $\sigma^2_{y.x}$ the quantity:

$$s^2_{y.x} = \{\Sigma Y^2 - (\Sigma XY)/(\Sigma X^2)\}/(n-1)$$

An estimate of the proportion of the variance of Y, as explained by this model, is the coefficient of determination d_{yx}:

$$d_{yx} = 1-(s^2_{y.x}/s^2_y).$$

It is interesting to note that, in general, $d_{yx} \neq d_{yx}$; i.e., d is not a symmetrical measure such as the squared correlation coefficient r^2. Furthermore, because the lower bound of d is not equal to zero, d cannot always be interpreted as a proportion. That d is nonetheless chosen instead of r^2 is because r^2 should be interpreted in terms of explained variance by the model $Y=\beta X+\alpha+\varepsilon$.

Aczél, J. (1966). *Lectures on functional equations and their application.* New York: Academic Press.

Adams, E.W., Fagot, R.F., & Robinson, R. (1965). A theory of appropriate statistics. *Psychometrika, 30,* 99-127.

Alpern, M., & Moeller, J. (1977). The red and green cone visual pigments of deuteranomalous trichromacy. *Journal of Physiology, 66,* 647-675.

Alpern, M., & Wake, T. (1977). Cone pigments in human deutan color vision defects. *Journal of Physiology, 266,* 595-612.

Barlow, H.B. (1981). Critical limiting factors in the design of the eye and the visual cortex. *Proceedings of the Royal Society of London, B 212,* 1-34.

Barlow, H.B. (1982). What causes trichromacy? A theoretical analysis using comb-filtered spectra. *Vision Research, 22,* 635-644.

Bastian, B. (1976). *Individual differences among the photopigments of protan observers.* Ph. D. Dissertation, Michigan: University of Michigan.

Bonet, D.G., & Bentler, P.M. (1983). Goodness-of-fit procedures for the evaluation and selection of log-linear models. *Psychological Bulletin, 93,* 149-166.

Boor, C. de. (1978). *A practical guide to splines.* New York: Springer Verlag.

Boynton, R.M. (1979). *Human color vision.* New York: Rhinehart & Winston.

Boynton, R.M., & Das, S.R. (1966). Visual adaptation: Increased efficiency

resulting from spectrally distributed mixtures of stimuli. *Science, 154,* 1581-1583.

Boynton, R.M., Das, S.R., & Gardiner, J. (1966). Interactions between visual photopic mechanisms revealed by mixing conditioning fields. *Journal of the Optical Society of America, 56,* 1775-1780.

Breton, M. (1979). A new model for predicting heterochromatic brightness matching. *Investigations of Ophthalmology & Visual Science, 19,* 207.

Brindley, G.S. (1960a). *Physiology of the retina and the visual pathway.* London: Arnold.

Brindley, G.S. (1960b). Two more visual theorems. *Quarterly Journal of Experimental Psychology, 12,* 110-112.

Buchsbaum, G. & Gottschalk, A. (1983). Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal Society of London, B 220,* 89-113.

Buckingham, E. (1914). On physically similar systems: Illustrations of the use of dimensional equations. *Physical Review, 4,* 345-376.

Burns, S.A., Smith, V.C., Pokorny, J., & Elsner, A.E. (1982). Brightness of equal-luminance lights. *Journal of the Optical Society of America, 72, 9,* 1225-1231.

Cameron, A.T. (1947). The taste sense and the relative sweetness of sugars and other sweet substances. *Scientific Report Series, No. 9.* Sugar Research Foundation.

Cavonius, C.R., & Hiltz, R. (1973). Brightness of isolated colored lights. *Journal of the Optical Society of America, 63, 8,* 913-920.

Cicerone, C., Krantz, D., & Larimer, J. (1975). Opponent-process additivity-III: Effect of moderate chromatic adaptation. *Vision Research, 15,*

1125-1135.

Cohen, J.D. (1975). Temporal independence of the Bezold-Brucke hue shift. *Vision Research, 15,* 341-351.

Comerford, J.P., & Kaiser, P.K. (1975). Luminous-efficiency functions determined by heterochromatic brightness matching. *Journal of the Optical Society of America, 65, 4,* 466-468.

Coombs, C.H., Raiffa, H., & R.M. Thrall. (1954). Some views on mathematical models and measurement theory. *Psychological Review, 61,* 132-144.

Cornsweet, T. (1962). The staircase method in psychophysics. *American Journal of Psychology, 75,* 485-491.

Cornsweet, T. (1978). The Bezold-Brucke effect and its complement, hue-constancy. In J.C. Armington, J. Krauskopf, & B.R. Wooten (Eds.), *Visual psychophysics and physiology.* New York: Academic Press.

Dagher, M., Cruz, A., & Plaza, L. (1958). Colour thresholds with monochromatic stimuli in the spectral region 530-630 nm. In *Visual Problems of Colour. I.* p. 319. N.P.L. Symposium No. 8. London: M.M.S.O.

DeValois, R.L., Abramov, I., & Jacobs, G.H. (1966). Analysis of response patterns of LGN cells. *Journal of the Optical Society of America, 56,* 966-977.

Dresler, A. (1953). The non-additivity of heterochromatic brightness. *Transactions of the Illumination Engineering Society, XVIII, 6,* 141-165.

Eisler, H. (1982). On the nature of subjective scales. *Journal of Psychology, 23,* 161-171.

Ekman, G., Eisler, H., & Künnapas, H. (1960). Brightness scales for monochromatic lights. *Scandinavian Journal of Psychology, 1,* 41-48.

Elzinga, C.H. (1984). Sensitivity in spectral structures. *Journal of*

Mathematical Psychology, 28, 4 421- 435.

Elzinga, C.H., & De Weert, Ch M M. (1978). Hue magnitude estimates as relative judgments. *Perception & Psychophysics, 23, 5,* 372-380.

Elzinga, C.H., & Ch M.M. de Weert (1984a) Invariance of direct brightness matching. *Journal of the Optical Society of America.* Submitted.

Elzinga, C.H., & De Weert, Ch M.M. (1984b). Nonlinear codes for the yellow/blue mechanism. *Vision Research, 24, 9* 911-922.

Estévez, O. (1979). *On the fundamental data-base of normal and dichromatic color vision.* Amsterdam (Ph.D. Dissertation)

Fletcher, H., & Munson, W A (1933). Loudness, its definition, measurement and calculation. *Journal of the American Acoustical Society, 5,* 82-108.

Frijters, J.E.R., & Oude Ophuis, P A.M (1983). The construction and prediction of psychophysical power functions for the sweetness of equiratio sugar mixtures. *Perception, 12,* 753-767

Guth, S , Donley, N., & Marocco, R. (1969). On luminance additivity and related topics. *Vision Research, 9,* 537-575.

Guth, S., & Lodge, H. (1973). Heterochromatic additivity, foveal spectral sensitivity and a new color model. *Journal of the Optical Society of America, 63,* 450-462.

Guth, S L., Massof, R W , & Benzschawel, T (1980). Vector model for normal and dichromatic color vision. *Journal of the Optical Society of America, 70,* 197-212.

Grum, F., & Becherer, R.J. (1979). *Optical Radiation Measurement. Vol. 1. Radiometry* New York· Academic Press.

Helmholtz, H.L.F. von (1896). *Handbuch der Physiologischen Optik.* Hamburg: Voss.

Heien, D.M. (1968). A note on log-linear regression. *American Statistical Association Journal, 63,* 1034-1038.

Hering, E. (1878). *Zur Lehre vom Lichtsinne.* Vienna: C. Gerolds Sohn.

Hooke, R., & Jeeves, T.A. (1961). Direct search solution of numerical and statistical problems. *Journal of the ACM, 8, 2,* 212-229.

Hurvich, L.M. (1977). Two decades of opponent processes. *Color 77,* International Color Association.

Hurvich, L.M., & Jameson, D. (1951). A psychophysical study of white. I. Neutral adaptation. *Journal of the Optical Soiety of America, 41,* 521-527.

Hurvich, L.M., & Jameson, D. (1957). An opponent process theory of color vision. *Psychological Review, 64,* 384-404.

Ikeda, M. (1983). Linearity law reexamined for flicker-photometry by the summation-index method. *Journal of the Optical Society of America, 73, 8,* 1055-1061.

Ikeda, M., & Ayama, M. (1980). Additivity of opponent chromatic valence. *Vision Research, 11,* 995-1000.

Ingling, C.R., Burns, S., & Drum, B.A. (1977). Desaturating blue increases only chromatic brightness. *Vision Research, 17,* 501-503.

Ingling, C.R., & Tsou, B.H.P. (1977a). Orthogonal combination of the three visual channels. *Vision Research, 17,* 1075-1082.

Ingling, C.R., & Tsou, B.H.P. (1977b). The spectral sensitivity of the opponent color channels. *Vision Research, 17,* 1083-1089.

Ingling, C.R., Tsou, B.H.P., Gast, T.J., Burns, S.A., Emerick, J.O., & Riesenberg, L. (1978). The achromatic channel - I. The nonlinearity of minimum-border and flicker matches. *Vision Research, 18,* 379-390.

Irtel, H. (1982). *Eine experimentelle und theoretische Analyse des biokularen Farbensehens. Thesis.* Regensburg.

Ives, H.E. (1912a). Studies in the photometry of lights of different colours. I. Spectral luminosity curves obtained by the equality of the brightness photometer and the flicker photometer under similar conditions. *Philosophical Magazine, 24,* 149-188.

Ives, H.E. (1912b). Studies in the photometry of lights of different colours. III. Distortions in the spectral luminosity curves produced by variations in the character of the comparison standard and of the photometric field. *Philosophical Magazine, 24,* 744-751.

Ives, H.E. (1912c). Studies in the photometry of lights of different colours. IV. The addition of luminosities of different colors. *Philosophical Magazine, 24,* 845-853.

Jacoby, S., Kowalik, J., & Pizzo, J. (1972). *Iterative methods for nonlinear optimization problems.* New Yersey: Prentice-Hall Inc.

James, F. & Roos, H. (1972) *Function minimization.* Proceedings of the 1972 CERN Computing and Data Processing School, Report CERN 72-21.

Jameson, D., & Hurvich, L.M. (1955). Some quantitative aspects of an opponent colors theory, I. Chromatic responses and spectral saturation. *Journal of the Optical Society of America, 45,* 522-546.

Johnson, N.L., & Kotz, S. (1970). *Continuous univariate distributions.* Vol. 2. New York: Houghton Mifflin Company.

Judd, D.B. (1951). Secr. Rep. "Colorimetry and artificial daylight". In: *Proceedings of the 12th Session of the C.I.E., Stockholm,* Vol. I, Technical CommiteeNo. 7, p.11.

Kohlrausch, A. (1935). Zur Photometrie farbiger Lichter. *Das Licht, 5,*

259-275.

König, A., & Dieterici, C. (1892). Die Grundempfindungen in normalen und anomalen Farbensystemen und ihre Intensitäts-Verteilung im Spectrum. *Zeitschrift fuer Physiologie der Sinnesorgane, 4,* 241-347.

Krantz, D.H. (1974). Measurement Theory and Qualitative Laws in Psychophysics. In Krantz, D.H., Luce, R., Atkinson, R.C., & Suppes, P. (Eds.), *Contemporary Developments in Psychophysics, Vol. II. Measurement, Psychophysics and Neural Information Processing, 160-199.* San Francisco: Freeman and Company.

Krantz, D.H. (1975a). Color measurement and color theory. I. Representation theorem for Grassmann structures. *Journal of Mathematical Psychology, 12,* 283-303.

Krantz, D.H. (1975b). Color measurement and color theory. II. Opponent colors theory. *Journal of Mathematical Psychology, 12,* 304-327.

Krantz, D.H., Luce, D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement. Vol. I.* New York: Academic Press.

Kruithof, A.A. (1953). Contributed discussion to: Dresler, A. The non-additivity of heterochromatic brightness. *Transactions of the Illumination Engineering Society, 18,* 141-165.

Larimer, J., Krantz, D.H., & Cicerone, C. (1974). Opponent process additivity I: Red/green equilibria. *Vision Research, 14* 1127-1140.

Larimer, J., Krantz, D.H., & Cicerone, C. (1975). Opponent Process additivity-II: Yellow/blue equilibria and nonlinear models. *Vision Research, 15,* 723-731.

Lemberger, F. (1908). Psychophysische Untersuchungen ueber den Geschmack von Zucker und Sacharin. *Archiven des gesammten Physiologie, 123,* 293-311.

Levelt, W.J.H., Riemersma, J.B., & Bunt, A.A. (1972). Binaural additivity of loudness. *British Journal of Mathematical and Statistical Psychology, 25,* 51-68

Luce, R.D. (1978). Dimensionally invariant laws correspond to meaningful qualitative relations. *Philosophy of Science, 45,* 1-16.

Luce, R.D., & Narens, L. (1983) Symmetry, Scale Types and Generalizations of Classical Physical Measurement *Journal of Mathematical Psychology, 27,* 44-85.

MacAdam, D.L. (1950). Loci of constant hue and brightness with various surrounding colors *Journal of the Optical Society of America, 40,* 589-595.

MacAdam, D.L. (1970). *Sources of Colour Science.* Michigan: MIT Press.

Mansfield, R.J.W. (1973). Brightness function· Effect of area and duration. *Journal of the Optical Society of America, 63, 8,* 913-920.

Mansfield, R.J W. (1976). Visual adaptation: retinal transduction, brightness and sensitivity. *Vision Research, 16,* 679-690.

Moeller, J.R. (1976). *Measuring the red/green quality of lights: A study relating the Jameson and Hurvich red/green cancellation valence to direct magnitude estimation of greenness.* Ph.D. dissertation. The University of Michigan.

Nagy, A.L. (1979). Unique hues are not invariant with brief stimulus durations. *Vision Research, 19,* 1427-1432.

Narens, L. (1981). A general theory of ratio scalability with remarks about the measurement-theoretic concept of meaningfulness. *Theory and Decision, 13,* 1-70.

Palacios, J. (1964). *Dimensional Analysis.* London: MacMillan & Co.

Palmer, G. (1777). *Theory of Colours and Vision.* London: Leacroft.

Pfanzagl, J. (1968). *Theory of Measurement.* Wurzberg-Wien: Physica-Verlag.

Pugh, E.N., & Larimer, J. (1980). Test of the site of yellow/blue hue cancellation and the site of the chromatic antagonism in the π_1 pathway. *Vision Research, 20,* 779-788.

Purdy, D.M. (1931). Spectral hue as a function of intensity. *American Journal of Psychology, 43,* 541-559.

Raaijmakers, J., & De Weert, C. (1975). Linear and nonlinear opponent color coding. *Perception & Psychophysics, 18,* 474-480.

Roberts, F.S. (1979). *Measurement Theory.* London: Addison-Wesley.

Robinson, D.W., & Dadson, R.S. (1956). A redetermination of the equal loudness relation for pure tones. *British Journal of Applied Physics, 7,* 166-181.

Rodieck, R.W. (1973). *The Vertebrate Retina, principles of structure and function.* San Francisco: Freeman & Company.

Romeskie, M. (1978). Chromatic opponent response functions of anomalous trichromats. *Vision Research, 18,* 1521-1532.

Roskam, E.E.Ch.I. (1974). *Unidimensional conjoint measurement (UNICON) for multi-facetted designs.* Report 74 MA 09, University of Nijmegen, Department of Mathematical Psychology.

Savoie, R.E. (1973). Bezold-Bruecke effect and visual nonlinearity. *Journal of the Optical Society of America, 63, 10,* 1253-1261

Schrödinger, E. (1920). Grundlinien einer Theorie der Farbenmetrik im Tagessehen. *Annalen der Physik, IV, 63,* 483-520.

Sirovich, L. (1977). On some mathematical techniques for the analysis of visual spectral sensitivity. *Biophysical Journal, 18,* 189-208.

Sirovich, L., & Abramov, I. (1977). Photopigments and pseudo-pigments. *Vision*

Research, 17, 5-16.

Smith, V.C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 700 nm. *Vision Research, 15,* 161-171.

Stevens, S.S. (1946). On the theory of scales of measurement. *Science, 103,* 677-680.

Stevens, S.S., & Davis, H. (1938). *Hearing, its psychology and physiology.* New York: John Wiley & Sons.

Stiles, W.S. (1946). A modified Helmholtz line-element in brightness-colour space. *The Proceedings of the Physical Society, 58,* 41-65.

Stiles, W.S. (1949). Increment thresholds & the mechanisms of colour vision. *Documenta Ophthalmologia, 3,* 138-163.

Stiles, W.S. (1955). The basic data of colour matching: 18th Thomas Young Oration. *Physical Society Yearbook,* 44-65.

Stiles, W.S. (1959). Colour Vision: The Approach through Increment Threshold Sensitivity. *Proceedings of the National Academy of Science, 45,* 100-114.

Suppes, P., & Zinnes, J.L. (1963). Basic Measurement Theory. In R.D. Luce, R.R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology. Vol. I.* New York: Wiley.

Szidarovszky, F., & Yakowitz, S. (1978). *Principles and procedures of numerical analysis.* New York: Plenum Press.

Teller, D.Y., & Pugh, E.N. (1983). Linking propositions in color vision. In J.D. Mollon, & L.T. Sharpe (Eds.), *Color vision, physiology and psychophysics.* London: Academic Press.

Tessier, M., & Bloittiau, M. (1951). Variations des characteristiques photometrique de l'oeil aux luminances photopiques. *Revue d'Optique, 30,*

309-322.

Thun, R.E. (1960). On dimensional analysis. *IBM Journal of Research & Development, 4,* 349-356.

Uttal, W.R. (1967). Evoked brain potentials: signs or codes? *Perspectives in Biology and Medicine, 10,* 627-639.

Uttal, W.R. (1973) *The psychobiology of sensory coding.* New York: Harper and Row.

Vos, J.J. (1978a). Colorimetric and photometric properties of a 2° fundamental observer. *Color Research & Application, 3, 3,* 125-128

Vos, J.J (1978b). *Tabulated characteristics of a proposed 2° fundamental observer.* Institute for perception TNO, Soesterberg, The Netherlands.

Vos, J.J., & Walraven, P.L. (1971). On the derivation of the foveal receptor primaries. *Vision Research, 11,* 799-818

Wagner, G., & Boynton, R.M. (1972). Comparison of four methods of heterochromatic photometry. *Journal of the Optical Society of America, 62, 12,* 1508-1515.

Walls, G.L. (1956). The G. Palmer story. *The Journal of the History of Medicine, 11,* 66-96.

Walraven, P.L. (1974). A closer look at the tritanopic convergence point. *Vision Research, 14,* 1339-1343

Wasserman, G.S. (1969). Heterochromatic additivity failure. *Psychological Review, 76, 2,* 221-223.

Wasserman, G S., & Gillman, C.B. (1970). Subadditivity and superadditivity of heterochromatic lights. *Psychological Review, 77, 4,* 338-342.

Wasserman, G.S (1978). *Color vision, an historical introduction.* New York: Wiley.

Werner, J S., & Wooten, B.R. (1979a). Opponent chromatic mechanisms: Relations to photopigments and hue naming. *Journal of the Optical Society of America, 69,* 422-434.

Werner, J.S., & Wooten, B.R. (1979b). Opponent chromatic response functions for an average observer. *Perception & Psychophysics, 25,* 371-374.

Wilson, B.C. (1964). *An experimental examination of the spectral luminosity concept.* Ph.D. dissertation. Ann Arbor, Michichigan: University Microfilms.

Winer, B.J. (1962). Statistical Principles in Experimental Design. New York: McGraw Hill Inc.

Wollenberg, A.L. van den, (1979). *The Rasch model and time-limit tests. An application and some theoretical contributions.* Nijmegen: Dissertation, Katholieke Universiteit.

Wright, W.D. (1946). *Researches on Normal and Defective Colour Vision* London: Henry Kimpton.

Wright, W.D. (1953). Contributed Discussion to : Dresler, A. The non-additivity of heterochromatic brightness. *Transactions of the Illumination Engineering Society, 18,* 141-165.

Wyszecki, G. & Stiles, W.S. (1967). *Color Science.* New York: Wiley.

Yaguchi, H., & Ikeda, M. (1983). Contribution of opponent-color channels to brightness. In J.D. Mollon, & L.T. Sharpe (Eds.), *Color vision, physiology and psychophysics.* London: Academic Press.

Yager, D., & Taylor, E. (1970). Experimental measures and a theoretical account of hue scaling as a function of luminance. *Perception & Psychophysics, 7,* 360-364.

Young, T. (1802). On the theory of light and colours. *Philosophical*

Transactions of the Royal Society of London, 92, 20-71.

Zaidi, Q. & J. Pokorny. (1983). Failures of metamerism for short wavelengths. *Journal of the Optical Society of America, 12,* 1902.

This dissertation is concerned with the measurement of sensation of color and brightness The major issue is the explicit development of a measurement procedure that is, in itself, a test of the meaningfulness of the data involved

Since the foundations of measurement and an associated concept of meaningfulness play such an important role in the material presented in this dissertation, Chapter 1 provides formal definitions of some elementary concepts from measurement theory.

Within the framework of measurement theory an empirical domain of interest is depicted as a set on which various relations and operations are defined that, together, formally denote the different phenomena that characterize the domain. Such a set, together with the defined relations is called a non-numerical relational structure. The proces of measurement then consists of the mapping of the non-numerical relational structure into a "similar" numerical relational structure Normally, a multitude of such mappings may in fact be used without altering the character of the correspondence between the numerical and the non-numerical relational structure Thus, a measurement structure consists of a triple the non-numerical relational structure, the numerical relational structure and the set of mappings from the former into the latter structure.

The most important theories on color and brightness, together with their database are then briefly discussed within the framework of measurement structures

Then, within this framework, a concept of meaningfulness is formulated that amounts to the notion that relations are called meaningful if and only if their content does not change under certain mappings of the domain of discourse onto itself.

From this point it is then argued that data are irrelevant, not if they are the result of subjective judgment, but if they are not meaningful in the sense indicated in Chapter 1.

Furthermore, Chapter 1 introduces the concept of a code in order to explain how properties of retinal receptors may be conceptually linked to properties of sensations.

It has since long been observed that the human observer is more responsive to lights of medium wavelength than to lights of shorter or longer wavelengths, just as is the case for pure tones of not too extreme frequencies. Such observations have led to procedures that "measure" the sensitivity of the observer to variation on various physical dimensions. Such procedures essentially amount to making the subject compare all wavelengths or frequencies with a standard wavelength or frequency and attenuating the power of one, such that the stimulus and the standard evoke equal strengths of the relevant sensation.

Chapter 2 provides a formal, measurement theoretical foundation for these procedures. It appears that the results of such a procedure are meaningful if, and only if, they do not depend upon the standard employed.

Furthermore, Chapter 2 formally describes the connection between, on the one hand, the measurement of sensitivity and, on the other hand, null-measurement. Null-measurement or, equivalently, cancellation consists in measuring an amount on one dimension that is required to restore balance or equilibrium of

a system that was perturbed by a manipulation on a different dimension. Numerous examples of this procedure may be found in physics and chemistry and it has been employed in color science to measure sensations of redness/greenness and yellowness/blueness.

The formal basis of the next two Chapters is so provided by Chapter 2.

In Chapter 3 the meaningfulness of brightness sensitivity is investigated along the lines that were, more generally and abstractly, drawn in Chapter 2. Many researchers in the field of vision held (and still appear to hold) the belief that sensations cannot be measured. For example, in 1953 W.D. Wright, when discussing a famous paper of Dresler (1953), held the opinion that "...since the sensation of brightness cannot be measured, the question of its additivity does not arise". However, the question had apparently arisen and indeed, as is demonstrated in Chapter 3, brightness can be measured and, using the properties of codes, a model for brightness sensitivity in terms of properties of retinal receptors is proposed that generates the kind of additivity failures that Wright and Dresler were discussing.

Chapter 4 discusses the status of a null-measurement procedure that employs sensations of yellowness and blueness. It appears that, according to the formal theory delineated in Chapter 2, sensitivity for yellowness and blueness cannot be meaningfully established since the results of the null-measurements are, according to the experiment discussed, definitely nonlinear. We proceed to discuss the various models that were proposed to account for this nonlinearity, then demonstrate that none of these models is capable of reproducing all relevant data and finally propose a new model that does account for these phenomena. This new model also predicts an, until then,

unobserved phenomenon· a tradeoff between wavelength and ıntensity for certaın lights that appear neither yellowısh nor bluish. From the second experiment treated ın Chapter 4 ıt appears that such an effect indeed occurs, thus corroboratıng the valıdıty of the new model.

Up to thıs point experıments have been dıscussed ın whıch the subjects task is to tune the intensity of a stimulus in such a way that the strength of the sensatıon ıt evokes matches to the strength of an sensatıon due to a dıfferent stımulus. In doing so the subject is not requıred to communicate such strengths to the experımenter. In Chapter 5 we employ the classıcal psychophysical paradigm of magnıtude estimatıon in order to study sensations of greenness, yellowness and brightness and their interdependence.
It is hypothesized that the magnitude estımates of yellowness and greenness of yellowısh green lıghts depend, on the one hand, upon the value of chromatic codes and, on the other hand, upon the magnitude estimate of the brightness of these same lights. The data of the experiment presented support this idea.
From the experiments reported it also appears that a kind of additivity failure of brightness, when assessed through sensitivity, also occurs with magnıtude estimates for brightness.

Dit proefschrift gaat over het meten van indrukken van kleur en helderheid. Het belangrijkste doel daarbij is de ontwikkeling van een meet-procedure die tegelijk een toets is voor de zinvolheid van de daarvoor vergaarde data.

Omdat de grondslagen van het meten en een daarmee samenhangend idee van zinvolheid zo'n belangrijke rol spelen in hetgeen in dit proefschrift wordt aangeboden, voorziet Hoofdstuk 1 in formele definities van enige elementaire begrippen uit de meettheorie.
Binnen het kader van meettheorie wordt een empirisch domein voorgesteld als een verzameling waarop verschillende relaties en operaties gedefinieerd zijn die, tezamen, de verschijnselen die dat domein karakterizeren formeel aanduiden. Zo'n verzameling, tezamen met de daarop gedefinieerde relaties, wordt een niet-numerieke relationele structuur genoemd. Het proces van het meten bestaat dan uit het afbeelden, of het construeren van een afbeelding, van de niet-numerieke relationele structuur in een "overeenkomstige" numerieke relationele structuur. Gewoonlijk kan een veelheid van zulke afbeeldingen gebruikt worden zonder dat het karakter van de correspondentie tussen niet-numerieke en numerieke relationele structuur verandert. Zo bestaat een meetstructuur dus uit een drietal: de niet-numerieke relationele structuur, de numerieke relationele structuur en de verzameling van afbeeldingen van de eerstgenoemde in de laatstgenoemde structuur.
Vervolgens worden de belangrijkste theorieen over kleur en helderheid en hun databasis kort besproken binnen het raamwerk van meet-structuren.

Binnen dat raamwerk wordt een formele definitie van zinvolheid gepresenteerd

die er op neerkomt dat relaties slechts zinvol zijn, wanneer hun inhoud niet verandert onder bepaalde afbeeldingen van het empirisch domein op zichzelf
Van hieruit wordt dan betoogd dat data irrelevant zijn, niet als ze het resultaat zijn van subjectief oordeel, maar als ze niet zinvol zijn in de betekenis zoals aangegeven in Hoofdstuk 1
Verder introduceert Hoofdstuk 1 het begrip code om uit te leggen hoe eigenschappen van retinale receptoren conceptueel verbonden zouden kunnen worden met eigenschappen van gewaarwordingen

Het is reeds lang bekend dat de menselijke waarnemer gevoeliger is voor golflengten uit het midden van het zichtbare spectrum dan voor lichten van een kortere of langere golflengte, zoals dat ook het geval is voor zuivere tonen van niet al te extreme frequentie Zulke observaties hebben geleid tot procedures waarmee de gevoeligheid van de waarnemer voor verschillende fysische dimensies kan worden "gemeten" Zulke procedures komen in essentie hierop neer dat van de proefpersoon gevraagd wordt alle golflengten of frequenties te vergelijken met een standaard-golflengte of een standaard-frequentie en het vermogen van de stimulus zodanig aan te passen dat de standaard en de stimulus een even sterke gewaarwording teweegbrengen
Hoofdstuk 2 levert een formele, meettheoretische grondslag voor zulke procedures Het blijkt dat de resultaten van zo'n procedure slechts zinvol zij wanneer ze onafhankelijk zijn van de standaard
Bovendien beschrijft Hoofdstuk 2 formeel het verband tussen, enerzijds, het meten van gevoeligheid en, anderzijds, nul-metingen Nul-metingen, ook wel cancellaties genoemd, bestaan uit het vaststellen van een hoeveelheid op een dimensie die nodig is om het evenwicht of equilibrium te herstellen van een systeem dat verstoord werd door een manipulatie op een andere dimensie In de

scheikunde en in de natuurkunde zijn de voorbeelden van deze procedure talrijk en ook is ze gebruikt in het kleurwaarnemingsonderzoek om de gewaarwordingen van rood/groenheid en geel/blauwheid te meten.
De formele grondslag van de twee volgende hoofdstukken is dan gelegd.

In Hoofdstuk 3 wordt de zinvolheid van helderheidsgevoeligheid onderzocht langs de wegen die, meer algemeen en abstract, in Hoofdstuk 2 werden gebaand. Veel onderzoekers op het gebied van visuele waarneming waren van mening (en lijken dat nog steeds te zijn) dat gewaarwordingen niet gemeten kunnen worden. In 1953 bijvoorbeeld, was W.D. Wright, een beroemd artikel van Dresler (1953) besprekend, van mening dat "...omdat de gewaarwording van helderheid niet gemeten kan worden, de vraag naar de optelbaarheid daarvan zich niet voordoet". Blijkbaar was de vraag echter opgekomen en, zoals in hoofdstuk 3 wordt aangetoond, kan de gewaarwording van helderheid inderdaad gemeten worden. Bovendien wordt, met gebruikmaking van eigenschappen van codes, een model, geformuleerd in termen van eigenschappen van retinale receptoren, voorgesteld dat de schendingen van optelbaarheid genereert waar Dresler en Wright over spraken.

Hoofdstuk 4 bespreekt de status van een nul-meting die gebruik maakt van de gewaarwordingen van geelheid en blauwheid. Het blijkt dat, volgens de formele theorie van Hoofdstuk 2, de gevoeligheid voor geelheid en blauwheid daarmee niet zinvol kan worden vastgesteld omdat de resultaten van die nulmetingen, blijkens het hier besproken experiment, niet lineair zijn. We bespreken vervolgens de verschillende modellen die werden voorgesteld om die non-lineariteit te verklaren, tonen dan aan dat geen van die modellen in staat is alle relevante data te genereren en stellen tenslotte een nieuw model voor dat

die verschijnselen wel verklaart. Dit nieuwe model voorspelt ook een nog niet eerder waargenomen verschijnsel: een afhankelijkheid tussen golflengte en intensiteit van bepaalde lichten die als noch geel, noch blauw worden waargenomen. Uit het tweede, in Hoofdstuk 4 behandelde, experiment blijkt dat zo'n effect inderdaad optreedt, waarmee de geldigheid van dat nieuwe model wordt bevestigd.

Tot zover zijn steeds experimenten besproken waarin het de taak van de proefpersoon is de intensiteit van een stimulus zodanig aan te passen dat de sterkte van de gewaarwording die zij oproept overeenkomt met de sterkte van eenzelfde gewaarwording ten gevolge van een andere stimulus. Daarbij wordt de proefpersoon niet gevraagd de proefleider van de sterkte van zijn gewaarwording te rapporteren.

In Hoofdstuk 5 gebruiken we het klassieke psychofysische paradigma van grootteschattingen om gewaarwordingen van groenheid, geelheid en helderheid te bestuderen. Verondersteld wordt dan dat de grootteschattingen van geelheid en groenheid van geel-groene lichten afhangt van, enerzijds, de waarde van codes voor chromaticiteit en, anderzijds, van de helderheidsgewaarwording tengevolge van die zelfde lichten. De data uit het hier besproken experiment blijken dit idee te ondersteunen.

Uit de gerapporteerde experimenten blijkt eveneens dat een type schending van additiviteit van helderheid zich voordoet dat gelijk is aan het type dat ook bij helderheidsgevoeligheid een rol speelt.

CURRICULUM

De schrijver van dit proefschrift werd geboren te Hilversum op 18 januari 1950. Hij behaalde het diploma 3-jr. HBS aan het Jordan Lyceum te Zeist, doorliep daarna de Christelijke Hogere Landbouwschool te Dronten en behaalde aldaar het diploma in 1970. Vanaf 1970 studeerde hij psychologie, met als hoofdrichting mathematische psychologie, aan de Katholieke Universiteit te Nijmegen en deed er zijn doctoraal-examen in 1977. Vanaf februari 1978 tot jananuari 1981 was hij als wetenschappelijk medewerker van ZWO en vanaf september 1981 van de Katholieke Universiteit te Nijmegen, verbonden aan de vakgroep Mathematische Psychologie.

Stellingen

1. Spectrale gevoeligheidsfuncties zijn slechts dan zinvol als de eraan ten grondslag liggende equivalentie invariant is.
(Dit proefschrift.)

2. Directe heterochromatische helderheidsvergelijkingen zijn invariant. Het model van Yaguchi en Ikeda (1983) voor deze vergelijkingen is niet homogeen en kan daarom niet adequaat zijn.
(Dit proefschrift.)

   Yaguchi, H., & Ikeda, M. (1983). Subadditivity and superadditivity in heterochromatic brightness matching. *Journal of the Optical Society of America, 23, 12,* 1711-1718.

3. Het theoretisch begrip "luminantie" (Wyszecki en Stiles, 1967) heeft vooralsnog geen empirische referent.
(Dit proefschrift.)

   Wyszecki, G., & Stiles, W.S. (1967). *Color Science, Concepts and methods, quantitative data and formulas.* New York: Wiley & Sons, Inc.

4. Een quantitatieve theoretische verklaring voor kleurindrukken behoort niet gesteld te worden in termen van opponente kleurmechanismen.
(Dit proefschrift.)

5. Het fundament van de dimensionele analyse is een geformaliseerde opvatting over de zinvolheid van relaties.

6. Ten onrechte suggereren Lehner en Noma (1983) dat het ORDMET-algorithme een bruikbare analyse-methode is voor data die verkregen zijn uit een proefopzet met meer facetten.

   Lehner, P.E., & Noma, E. (1980). A new solution to the problem of finding all numerical solutions to ordered metric structures. *Psychometrika, 45, 1,* 135-137.

7. De constructie van de maximin-r^2 vector in een convex polyhedron is equivalent met de constructie van het centrum van de hyperbol met de kleinste inhoud die bepaald wordt door de genormaliseerde vertices van dat polyhedron.

8. Vele van de ideeën die ten grondslag liggen aan het co-counselen zijn strijdig met de principes die emancipatorische bewegingen trachten te realiseren en missen bovendien iedere wetenschappelijke fundering. Het bedrijven of propageren van co-counselen is daarom politiek en ethisch onjuist.

9. De door Thomas (1981) voorgestelde schatters van de vermenigvuldiger in psychofysische machtsfuncties met een log-normaal verdeelde fouten-component zijn niet de meest optimale schatters.

Thomas, H. (1981). Estimation in the power law. *Psychometrika, 46,* 29-34.

Elzinga, C.H. (1985). A note on estimation in the power law. *Perception & Psychophysics.* (In druk.)

10. Bevorderlijk voor de kwaliteit van het onderzoek in de psychologie, in het bijzonder ook in de mathematische psychologie, zou zijn dat mathematisch psychologen zouden werken binnen een vakgroep waarin de inhoudelijke aspecten van hun studie-object een centrale plaats innemen.